EXPLORING THE USABILITY OF AUGMENTED REALITY INTERACTION TECHNIQUES DURING CHILDREN'S EARLY ELEMENTARY-SCHOOL YEARS

A Dissertation
Presented to
The Academic Faculty

By

Iulian Radu

In Partial Fulfillment
Of the Requirements for the Degree
Doctor of Philosophy in Human-Centered Computing

Georgia Institute of Technology

December 2016

EXPLORING THE USABILITY OF AUGMENTED REALITY INTERACTION TECHNIQUES DURING CHILDREN'S EARLY ELEMENTARY-SCHOOL YEARS

Approved by:

Dr. Blair MacIntyre, Advisor
School of Interactive Computing
Georgia Institute of Technology

Dr. Stella Lourenco
School of Psychology
Emory University

Dr. Ashok Goel
School of Interactive Computing
Georgia Institute of Technology

Dr. Alissa N. Antle
School of Interactive Arts and Technology
Simon Fraser University

Dr. Ellen Yi-Luen Do
School of Interactive Computing
Georgia Institute of Technology

Date Approved: August 17, 2016

ACKNOWLEDGEMENTS

I would like to express infinite gratitude for my parents, Simona Radu and Adrian Radu, without whom I would not have been able to have this personal growth and amazing life experience, and whom I know are watching this moment with pride.

I thank the following people for providing me with vital emotional support, reminding me that there is life outside of the ivory tower, and sharing the beauties of their spirits with mine: Ieva Mikolaviciute, Ogechi Nnadi, Jason Tsukahara, Ben Perrodin, Radu Pop, Aaron Smith, Lea Young, Cristian Csohan, Dana Ionescu, Anne Franklin, Matt Sniatynski, Julie Fabelhaft-Chan, Bogdan & Ioana Morariu, Clara & Gelu Danes, Sorina & Lucian Pop, Sorina & Silviu Pop, Rosa Arriaga, Ruby Zheng, Yan Xu, Michael Hewner, Ceara Byrne, Manasvi Lalwani, Craig Tashman.

I thank my advisor Blair MacIntyre for putting up with my adventurous research style, and the other members of my thesis committees, who provided their time and wisdom to advise, grill and strengthen this research: Alissa Antle, Ashok Goel, Ellen Do, Stella Lourenco, Amy Bruckman, Nancy Nersessian, Gregory Abowd. I thank my other collaborators who have encouraged wild ideas and shaped academic publications: Keith Bujak, Lana Yarosh, Evan Barba, Maribeth Gandy, Saga HaiTao Song, PBS KIDS, WestEd, WNET. Thank you to the hard working undergrads who have strongly influenced the research studies and memorized countless versions of video coding schemes: Simina Avram, Katie Guzdial, Larissa Wildenburg, Meera Nathan, Taeho Kim, Raveena Singh. I also thank the following people and organizations for volunteering their valuable time and advice to make me a better researcher: Sahithi Bonala, Andrea Lau, Jen Rodriguez, Emory Spatial Cognition Lab, Richard Catrambone, Georgia Tech's Problem Solving and Educational Technology Lab, Gheric Speigner, Taeheon Kim, Jessica Celestine, and The Children's School. And thank you to the anonymous child volunteers, parents, teachers, and reviewers who have helped in the production of the games, studies, and publications that were generated through this research.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

SUMMARY

Augmented reality (AR) has been shown to have measurable benefits in enriching children's lives, by advancing education through in-situ 3D visualizations, providing entertainment through whole-body interaction, and enhancing physical & cognitive rehabilitation through motivational engagement. Although such experiences were typically confined to desktop computers, the increasing popularity of mobile devices is expected to make AR accessible to large amount of children. In order to realize these benefits, technology designers need to create experiences that are usable by children. Handheld AR interfaces are different from more traditional interfaces, by being small portable windows into physical spaces augmented with digital content, and their use may require users to employ more complex motor and cognitive skills than compared to traditional interfaces. Due to the novelty of handheld AR technology, there are no standard interaction techniques for handheld AR, and little is known about children's ability to use these interfaces.

In the current research, I address the following questions: How does children’s age relate to performance and usability issues in handheld-AR? How do different handheld-AR interaction techniques compare, in terms of performance and usability issues encountered by children? And, what types of usability issues are experienced by children in handheld-AR?

In order to address these questions, I first constructed several commercial and prototype educational AR games for young children and studied their educational potential, as well as children's ability to use these games. I contributed analyses of how augmented reality can be applied in educational contexts. Further, I generated a usability framework that organizes the usability issues observed in my studies and in existing literature on AR systems for children, discusses relationships between developmental psychology literature and children's AR usability, and provides guidelines for designing AR for children. Finally, I performed a systematic study of children 5-10 years old using handheld augmented reality, as they played a smartphone-based AR game using two AR interaction techniques (finger-based vs. crosshair-based selection) under two movement conditions (tunnels vs. no tunnels). Children's performance and usability problems

were analyzed through quantitative and qualitative methods. This research identifies complex relationships between usability metrics and children's age across the elementary-school years (e.g., significant changes occurring around 7 years, in children's ability select items quickly while reorienting their body in a 3D space; significant increase in poor postures in older children; age-invariant frequency of phone dropping; etc.). This research also identifies a variety of usability issues encountered by children of different ages (e.g., the detrimental effects that previous exposure to non-AR technology has on children's ability to work with AR tracking technology; or, the variety of ways in which children lose tracking while playing AR games). The research identifies links to cognitive and physical developmental skills that underlie AR performance (e.g., crosshair-based selection employs skills related to visuomotor precision and spatial relations, while finger-based selection employs skills used in block building activities). Furthermore, gender differences in technology exposure are identified, along with effects of practice during short gameplay sessions. This research concludes with a set of guidelines for designing handheld AR technology for children in the 5-10 year old range, along with a set of directions for future work, involving applications of child usability to upcoming AR technologies, and improvements to the research methods in studying AR for children.

CHAPTER 1

INTRODUCTION

There are many potential benefits which augmented reality (AR) technology can bring to children's lives, such as enhanced entertainment through whole-body interaction (Billinghurst 2002, De Lisi and Wolford 2002), advancing education through in-situ interactive visualizations (Shelton and Hedley 2003, Kerawalla, Luckin et al. 2006), and improving rehabilitation and skill development through physical manipulation (Merians, Jack et al. 2002, Tang, Owen et al. 2003). In the domain of education, augmented reality has been shown to have measurable benefits over traditional approaches (see (Radu 2012) for a review). It has been argued that AR may replace physical manipulatives in the elementary school classroom (Bujak, Radu et al. 2013), and educational organizations such as PBS and Sesame Workshop have begun to investigate the use of AR for elementary children's education (Public Broadcasting Service 2011, Sesame Workshop 2013).

To achieve these benefits, augmented reality experiences need to be appropriately designed for young children. Previous research on technology design for children has stressed the importance of age-appropriate design (Baumgarten 2003, Hourcade 2008, Gelderblom 2009, Bekker and Antle 2011). However, there is a lack of systematic understanding of how to design AR experiences for children in the augmented reality design community (Radu and MacIntyre 2012). Existing guidelines developed in non-AR media, such as for desktop PC environments, are likely to have limited applicability to AR, due to the differences between AR and non-AR media (Radu 2012, Bujak, Radu et al. 2013, Radu 2014). Specifically, augmented reality introduces virtual elements into a user's physical environments, thus interacting with AR likely causes higher demands on motor skills (since users interact by moving in a 3D space), higher demands on spatial cognition (since users must understand spatial relationships between virtual and physical objects located in 3D space), higher demands on attention control (since users must distinguish and attend to two versions of the physical world: the display showing the augmented reality, and the non-augmented real space where the user interacts), and may require a different conceptualization of the digital experience (since users must conceptually understand that parts of what they observe in the physical world is computer-generated content).

Augmented-reality interfaces are therefore significantly different than traditional interfaces, and it is important to develop our understanding of children's reactions to this technology. I am currently working with PBS on designing augmented reality applications for young children, and I have experienced firsthand the difficulties caused by this lack of knowledge. Although PBS has access to a variety of experts knowledgeable about children's entertainment and education with traditional media, and although previous research shows that AR can be a valuable technology for education, technology designers are unclear how to design AR experiences for children and unclear about whether children can use the handheld-AR applications that PBS envisions. On one hand, this problem is due to the fact that there exist no specific usability guidelines for designing AR applications for children. There are case studies researching AR applications for children on platforms such as webcams, projectors and head-mounted displays, but no specific guidelines have yet been generated. Furthermore, almost no research publications exist about children's experiences on handheld platforms such as iPhones and Android phones, which have the potential to bring in-situ educational experiences to a massive population of children. On the other hand, the problem is intensified by the fact that young children experience significant developmental changes during a few years of life, and these developments may impact their ability to use this novel technology; thus, it is unclear (and often unlikely) that young children will be able to use AR designs suitable for older children or adults.

In my research, I am interested in extending our understanding of AR usability for children, by specifically investigating the abilities of young children (aged 5-10) to use handheld-AR interactions. Augmented-reality can be realized on a variety of platforms, such as desktop computers, handheld smartphones, head-mounted displays, yet I am specifically interested in handheld-based augmented reality since I foresee that the continuing increase in popularity of smartphone devices will make AR accessible to large amount of children. Because of the accessibility of handheld-AR experiences, the medium will become a significant channel for leveraging the education impacts of augmented reality for elementary-school students. Very few handheld-AR interfaces for children have been empirically studied, and I believe it is important to understand how children react to handheld-AR since it is different than other AR platforms, due to its high portability and use of small display sizes. The results of my research will be therefore be specifically applicable to handheld-AR; however, it is possible that results may be generalizable to other kinds of AR experiences. In the text of this dissertation, I will note when the discussion applies to handheld-AR or to AR in general.

Furthermore, I am focusing specifically on usability issues that are related to interaction mechanics. For example, in an AR game that requires children to select virtual items located on a marker, I am interested to know whether children have trouble understanding that virtual items have locations in physical space, whether children will be able to physically move their device close to the virtual items, whether children will frequently lose tracking while they play the game, and so on.

1.1 RESEARCH QUESTIONS, HYPOTHESES AND CONTRIBUTIONS

The purpose of my research is to investigate young children's ability to use different interaction techniques for handheld augmented reality. In my studies, quantitative instruments will be used to measure the relationship between child performance and interaction technique type, as well as the relationship between performance and child age. At the same time, usability issues will be explored through qualitative observations of children interacting with various augmented reality applications. I intend to investigate the following research questions:

RQ1: How does children's age relate to performance and usability issues in handheld-AR?

RQ2: How do different handheld-AR interaction techniques compare, in terms of performance and usability issues encountered by children?

RQ3: What types of usability issues are experienced by children in handheld-AR?

Table 1.1.1 lists each of these primary questions, their sub-questions and hypotheses.

Table 1.1.1. Research questions and hypotheses.

RQ1: How does children's age relate to performance and usability issues in handheld-AR?
RQ1-1: Does speed of selection differ between age groups? ***Hypothesis***: Younger children will be slower at performing selection tasks. **RQ1-2: Does selection accuracy differ between age groups?** ***Hypothesis***: Younger children will be less accurate at performing selection tasks. **RQ1-3: Does accuracy for AR tracking differ between age groups?** ***Hypothesis***: Younger children will have a higher frequency of AR tracking losses. **RQ1-4: Does speed of AR tracking recovery differ between age groups?** ***Hypothesis****:* Younger children will be slower at recovering the AR tracking. **RQ1-5: Does children's self-reported fun, ease-of-use, and comfort change with age?** ***Hypothesis1****:* As age increases, there will be an increase in self reported ease-of-use. ***Hypothesis2****:* As age increases, there will be an increase in self reported comfort. ***Hypothesis3****:* It is unclear if fun would change across age groups. **RQ1-6: Do usability issues differ between age groups?** ***Hypothesis***: Younger children will experience higher number of usability issues.

Table 1.1.1. (Continued)

RQ2: How do different handheld-AR interaction techniques compare, in terms of performance and usability issues encountered by children?
RQ2-1: Does speed of selection differ between interaction techniques? ***Hypothesis1***: Interaction techniques that involve independent hand movements will lead to lower speed. ***Hypothesis2***: Interaction techniques that involve whole-body movement will lead to lower speed. **RQ2-2: Does selection accuracy differ between interaction techniques?** ***Hypothesis1****:* Interaction techniques that involve independent hand movements will lead to lower accuracy. ***Hypothesis2****:* Interaction techniques that involve whole-body movement will lead to lower accuracy. **RQ2-3: Does accuracy for AR tracking differ between interaction techniques?** ***Hypothesis1***: Interaction techniques that involve independent hand movements will lead to higher frequency of tracking losses. ***Hypothesis2****:* Interaction techniques that involve whole-body movement will lead to higher frequency of tracking losses. **RQ2-4: Does speed of AR tracking recovery differ between interaction techniques?** ***Hypothesis1****:* Interaction techniques that involve independent hand movements will lead to lower tracking recovery speed. ***Hypothesis2****:* Interaction techniques that involve whole-body movement will lead to lower tracking recovery speed. **RQ2-5: Does child development correlate with performance under different interaction techniques?** ***Hypothesis1***: Performance on interaction techniques that involve independent hand movements will be inversely correlated to tests of fine motor skills and physical manipulation. ***Hypothesis2***: Performance on interaction techniques that involve whole-body movement will be inversely correlated to tests of fine motor skills, physical manipulation, and spatial skills.

Table 1.1.1. (Continued)

RQ2-6: Does children's self-reported fun, ease-of-use, and comfort change between interaction techniques? ***Hypothesis1****:* Interaction techniques requiring independent hand movements or whole-body movements will yield less self-reported ease-of-use. ***Hypothesis2****:* Interaction techniques requiring independent hand movements or whole-body movements will yield less self-reported comfort. ***Hypothesis3****:* It is unclear if fun would change across interaction techniques. **RQ2-7: Do usability issues differ between interaction techniques?** ***Hypothesis***: Interaction techniques that involve independent hand movements will lead to more usability issues. ***Hypothesis***: Interaction techniques that involve movement difficulty will lead to more usability issues.
RQ3: What types of usability issues are experienced by children in handheld-AR?
Hypothesis: Children will experience problems that can be linked to developing areas of physical and cognitive skills

I will now discuss the main research questions and their associated contributions. For a summary of the methods and results associated to each research question, please refer to Appendix A.

1.1.1 RQ1: How does children's age impact performance and usability issues in handheld-AR interactions?

Age is an important factor in application design for children, especially critical for the design of interaction techniques in educational applications. In the typical school environment, children are expected to master specific concepts at specific grade levels (Common Core Standards Initiative 2012); therefore, in order to create effective educational technologies, designers must understand usability of children at targeted grade levels (Gelderblom 2009). In the augmented reality community, no research has systematically studied differences between children of different ages. The community does not understand what kinds of usability issues are encountered by children of different ages, there are no systematic comparisons of usability between different ages, and there is no data about what performance can be expected at different ages. My research investigates these issues as they apply to several interaction techniques.

To answer the research question, I have designed a series of handheld augmented reality games and exposed them to children of different ages. Through informal and experimental studies, I have gathered data on children of early elementary school age, specifically focusing on the 5-10 year old range. This age range has been chosen because it spans across significant changes identified by developmental psychology, and because it represents a fruitful area for AR educational applications, appealing to elementary-school children that are moving from playing with concrete toys, to understanding abstract concepts.

A systematic comparison between age groups has been performed through my main experimental study (Section 4). Qualitative data was used to determine the type and severity of usability problems experienced by children as they play each game. Quantitative performance recorded children's performance on a variety of metrics, including speed and accuracy during gameplay. The qualitative and quantitative data was aggregated and compared between age groups, to determine how age impacts children's reactions to AR. In the study, I have also investigated other variables such as gender, previous experience with technology, in order to determine relationships to AR usability. The contribution from this experiment is a systematic comparison between AR interactions of three age groups: children 5-6, 7-8, and 9-10 years old. The analysis indicates a developmental shift between ages 5-6 and 7-8, whereby children 5-6 show significantly poorer performance on the majority of AR metrics when compared to both 7-8 and 9-10 year olds; however, children 7-8 and 9-10 years old are not significantly different on most of the metrics analyzed. Furthermore, younger children experience a greater number and higher severity of usability issues.

1.1.2 RQ2: How do different handheld-AR interaction techniques compare, in terms of performance and usability issues encountered by children?

Designers of children's applications have a wide repertoire of interaction techniques to choose from, when selecting the input channel from the user to the application. In PC based applications, the input devices of keyboard and mouse have been well established, and there have been many empirical studies to establish child-appropriate interaction techniques (Inkpen 2001, Hourcade 2008, Hourcade, Perry et al. 2008). In contrast, the medium of handheld augmented reality is a relatively new phenomenon, and very few studies have empirically investigated handheld-AR usability for children (Radu and MacIntyre 2012). No systematic studies have been performed to compare children's reactions to different handheld-AR interaction techniques. Thus, designers of children's handheld-AR do not understand what interaction techniques are appropriate for young children, nor is it known what performance and usability issues children encounter as they use a specific interaction technique.

To answer this research question, I have exposed children to different interaction techniques in various handheld-AR game experiences. The interaction techniques varied between children simply holding the device with two hands and moving their body in space; to needing to touch on the screen with at least one finger; to aiming with a crosshair in the center of the screen; or needing to hold the device with one hand while another hand is manipulating a paddle or other object. During my main experimental study I have studied two popular interaction techniques – finger selection and crosshair selection, under two whole-body movement conditions where the game required movement around the gameboard, or not. Quantitative data was collected to determine user performance between interaction techniques, and qualitative analysis was used to determine the types and severity of usability issues associated with teach technique. These studies indicate that finger and crosshair are different only on a few metrics; in contrast, the whole-body movement conditions yield a variety of differences in terms of performance and usability issues. Existing literature and my own previous work indicate that child performance and usability issues may be correlated with the motor and spatial complexity of interaction techniques; thus, I have also investigated potential correlations between usability and standardized metrics of child development. The analysis shows that the two interaction techniques of finger vs. crosshair selection use different developmental skills, related to block construction and visuomotor precision.

The contribution of this aspect of my research is a comparison between several handheld AR interaction techniques. The research shows differences in terms of performance, usability issues, children's preference, and gender. Further, the research shows differences between how children of different ages interact with different techniques.

1.1.3 RQ3: What types of usability issues are experienced by children in handheld-AR?

Developmental psychology tells us that children have certain capabilities and limitations, which are different than adults (Rosser 1994). These developing abilities mediate children's ability to use technology, and researchers have stressed the importance of considering developmental abilities when designing technology for children (Bruckman and Bandlow 2002, Wyeth 2003, Gelderblom 2009, Bekker and Antle 2011). Besides my work (Radu and MacIntyre 2012), no research in the augmented reality community has attempted to classify the different kinds of usability issues experienced by children, or to understand how children's ability to use AR may be affected by developmental skills.

To address this knowledge gap, I contribute a framework that (1) summarizes AR literature describing various usability issues encountered by children, (2) describes developmental factors which can account for the observed usability issues, and (3) provides guidelines for designing age-appropriate AR designs. The framework was built from my informal studies and literature reviews, and initially presented in (Radu and MacIntyre 2012) (described in Chapter 3.3 and Appendix 1). It was then refined based on the qualitative data from my main experimental study. The methodology for constructing the framework involved correlating developmental psychology literature with case studies from AR literature and with my own pilot study observations. Subsequently, a coding scheme was constructed based on the usability framework and videos of children performing AR tasks (presented in Chapter 3.3). Finally, the coding scheme was used to identify usability issues on children 5-10 years old, leading to refinement of the usability framework and associated guidelines.

1.1.4 Other Contributions

Based on this research, I contribute a set of guidelines for designing child-friendly AR applications. I summarize the research findings by discussing specific AR designs, identifying the kinds of usability issues encountered by children in reaction to specific designs, discussing the suitability of existing interaction techniques for children of different ages, and presenting ideas for modifying the interaction techniques to make them more child-friendly.

The experimental game used in the research studies will be a contribution to the research community, and will be uploaded to a website. The experimental game used in my main study can function as a software testbed for testing different kinds of interaction techniques in a standardized environment, similar to the VR-based software testbed presented by (Bowman and Hodges 1999). This software will be made available to other researchers who wish to perform systematic comparisons between handheld AR interaction techniques.

The coding scheme used for detecting usability issues in my research studies will be published or made available on a website. This coding scheme will be useful to the AR community as an instrument for detecting usability issues, and comparing usability issues between experimental settings.

Finally, through the research I have generated various prototype games for children. These games will be made available on a website, and will serve as design inspiration for producers of children's AR applications.

CHAPTER 2

RELATED WORK

The related work is composed of three different sections. First, I discuss differences in children of different ages, based on research from developmental psychology and human computer interaction, and reflecting on its impact to children's handheld augmented reality design. Then, I discuss the notion of interaction techniques, focusing on different design aspects that influence user's performance and usability issues, and how these may influence interaction design for children's handheld-AR. Finally, I discuss work from augmented reality, describing the various interaction techniques used in applications for children, and the various usability issues reported by existing research.

2.1 DEVELOPMENTS ACROSS CHILDHOOD

In the following discussion, I will focus on children's development in the areas of general cognition, understanding of space, and ability to perform motor actions. I have selected these areas of development because I believe they significantly influence children's ability to use handheld-AR applications. I will now discuss each area by first presenting research from developmental psychology, then discussing its implications for handheld augmented reality.

2.1.1 Cognitive Development

Jean Piaget is arguably the most influential developmental psychologist of the 20th century. He is known for his extensive study of children's development, most notably for his theory of developmental stages, which posits that children's cognitive development occurs in sequential stages, each with its defining characteristics. Piaget defined 4 stages of child development: sensorimotor stage (birth – year 2), preoperational stage (years 2-7), concrete operational stage (years 7-11) and formal operations stage (age 11 and onward) (Piaget 1970). It is worth noting that researchers have criticized the specific age boundaries proposed by Piaget, and it is not expected that children will show a radical change in skill at the boundary age of 7. Research criticizing Piaget has shown high variability in children, specifically that children within the same age group do not all possess the same cognitive capabilities; and, depending on experimental conditions, children can show some cognitive competencies at an earlier age than Piaget's stages predict (Flavell, Beilin et al. 1992, Rosser 1994, Fischer and Immordino-Yang 2002, Thornton 2002, Kesselring and Müller 2011). Nevertheless, Piaget's stages can be useful for illustrating the general changes that occur in children's cognition and for providing a set of initial predictions about the variety of usability issues that children may encounter with augmented reality.

According to Piaget, significant changes occur when children move from the preoperational stage to the concrete operational stages. During the preoperational stage, children have learned to use language and symbols, they actively use their imagination, and engage in pretend play (Rosser 1994). Children's view of the world at this stage is egocentric, and they have difficulty understanding other people's points of view (Piaget 1970, Rosser 1994). This egocentric behavior disappears during the concrete operations stage, and children begin to think strategically about other people's actions as well as their own (Rosser 1994). Children can typically pay attention to only one item at a time during the preoperational stage (Piaget 1970), but attention skills develop to attending toward multiple items during the concrete operational stage. In terms of memory at the preoperational stage, children can remember a limited number of items (at 5 years old, children can remember 4-5 items), have difficulty remembering items in reverse, and working with hierarchies (Dempster 1981, Hourcade 2008). These memory skills develop during the concrete operational stage, as children can remember more items (roughly 6 items by age 9), and are able to remember items in reverse order and work with hierarchies (Dempster 1981, Hourcade 2008). Reasoning skills are based on perception during the preoperational stage: children have trouble with logical inference, and they tend to focus on surface features of situations, as exemplified by children's failure of the water conservation test (Rosser 1994). During the concrete operations stage, logical thinking skills develop and children learn to reason about invisible aspects of situations, and can pass the conservation task (Rosser 1994). Although logical thinking skills improve during the concrete operations stage, children are still influenced by their imaginations and by evidence from their senses, rather than logic (Piaget 1970, Rosser 1994).

These observations indicate that young children may experience various usability issues while interacting with handheld augmented reality. The developing logic skills may lead young children to be unable to understand that the objects observed on the screen are artificially generated inside the handheld device, and do not exist in the real world. Alternatively, young children's tendency to focus on concrete aspects of a situation may lead children to focus on the concrete handheld device, and children may not understand that the depicted virtual objects are supposed to be associated with the physical space in front of the player. Further, the attention and memory limitations may make children unable to attend to important aspects of the game (e.g., not noticing when virtual objects appear or disappear, or not remembering the previous states of game objects) or being unable to attend to the physical game environment (e.g., bumping into physical objects while playing the game, or not noticing when virtual objects change their physical location). Finally, children may be unable to build a correct mental model of the technology or the interaction mechanic, leading to lack of understanding of how to perform game interactions, or inability to understand how to recover from loss of AR tracking.

2.1.2 Spatial Development

Regarding the development of spatial cognition, significant development occurs around the age of 7 years of age. Piaget believed that until age 7, children encode spatial relationships using egocentric frames of reference, remembering objects as relative to one's body rather than relative to other objects in the world (Rosser 1994). Further, Piaget indicated that preoperational children do not use metric encoding to remember distances between themselves and other objects, but rather use a topological encoding method, whereby children remember spaces in terms of the relative order of objects (Rosser 1994). Only around 9-10 years old do children appear to encode spatial relationships in terms of metric distance, and in allocentric (object-centered) encoding (Rosser 1994). Research into the spatial visualization abilities of children indicates that children have trouble with line-of-sight predictions until around 6-7 years old, and that children younger than 6 years old have trouble imagining how items in a space are arranged when seen from another person's perspective (Piaget 1970, Rosser 1994, Payne and Isaacs 2002). Additionally, before 6 years old children have trouble imagining how a space would look in a mirror (Rosser 1994). Further, until roughly 6 years old, young children have trouble performing mental rotation, and have trouble predicting the intermediate stages of motion, for instance when being asked to infer between two snapshots of a falling stick (Rosser 1994). Small vs. large spaces are also an area of difference between younger and older children, and neuroscience research shows that processing of spatial information activates two different brain areas depending on the spatial scale (Bullens, Nardini et al. 2010). In studying children's use of landmarks, researchers have indicated that children younger than 7 bias their encodings based on landmarks close to the target, rather than landmarks in the broader environment (Bullens, Nardini et al. 2010, Vasilyeva and Lourenco 2010); in contrast, older children and adults use landmarks in the broader environment to remember locations in a space (Bullens, Nardini et al. 2010).

Gender differences in spatial cognition have been detected in children as young as 3 months old. Infants of different genders show varied looking times when presented with novel stimuli during a rotation tasks; it is unclear, however, if one gender has an advantage on these tasks, or whether children of different genders process spatial representations differently (Quinn and Liben 2008, Levine, Foley et al. 2016). Males do show better performance on tasks of 2D mental rotation and transformation by the age of 4½ years (Levine, Huttenlocher et al. 1999). Gender differences have mostly been studied for teens and adults, with males exhibiting higher performance for visuospatial skills (Crucian and Berenbaum 1998, Geary and DeSoto 2011), and females exhibiting higher performance for spatial location memory (Voyer, Postma et al. 2007). Although studies of young children do show gender differences, it is unclear what factors cause these differences. Socioeconomic status (SES) has been shown to be linked with gender differences on spatial tasks. Research indicates that gender differences of 7-8 year olds may only occur in medium- and high-SES children, and this is possibly due to a combination of parental gender attitudes and increased access to spatially-engaging activities such as playing with Lego blocks, puzzles, videogames, and outdoor exploration (Levine, Vasilyeva et al. 2005). It is therefore possible that when children engage with augmented reality games, they may show gender effects (with better spatial performance in males), due to differences in spatial skills and/or due to children's socioeconomic status.

These findings indicate that children of different ages may differ in their understanding of augmented reality spaces. The observations related to children's egocentric encoding, and inability to visualize spatial transformations, indicate that children may have trouble understanding how to change their physical location to act upon the game space, and may have trouble predicting what virtual objects are visible by another game player. Not remembering or perceiving the actual distances between virtual objects may cause children inaccurately remember the spatial configuration of the AR space, and may cause children to be unable to physically interact with the virtual game elements, such as when the interaction requires children to reach into the space to touch virtual object locations, or when the game requires children to move around a space between virtual object locations. The differences between processing of large and small spaces, combined with children's reduced memory capacity, also presents the possibility that children may have trouble when interacting with a large virtual space through small handheld displays, since they may not have accurate memory for virtual objects that are not visible on the screen.

2.1.3 Motor Development

When children physically manipulate augmented reality, different skills come into play. Children must make use of both fine and gross motor skills, such as when touching virtual items on the handheld screen (a fine motor skill), or walking around the augmented space to reach far away items (a gross motor skill). Perceptual skills such as visual acuity and hand eye coordination will influence how well children can see and intercept the elements of the AR application. Finally, children may be required to employ bimanual coordination (e.g., coordinating hands independently), for instance when holding the device with one hand and touching the screen with the other.

The development of ability for fine motor movements is illustrated by children's ability to draw and write: By age 4-5, most children hold writing tools with the mature "tripod" finger grip, their precision has improved, and they are able to color within the lines of a coloring book; by 9-12, precision and hand-eye coordination matures, and children are able to write properly-spaced letters, and they can copy drawings by transferring from a grid (Hourcade 2008). In contrast, gross motor skills, which involve large muscles such as when walking or jumping, develop faster than fine motor skills. At ages 4-5, children begin to show mature forms of walking and running, and are able to run and kick a ball in a smooth movement (Payne and Isaacs 2002). By age 8, children show mastery of their movements by combining movements during running and dancing, such as the ability to change movement type, speed and direction on demand (Payne and Isaacs 2002).

Control of movement begins developing in infants, develops through early childhood, shows maturation during the years of middle childhood, and declines at older age (Yan, Thomas et al. 2000, Hourcade 2008). Young children's control of motor movements is not as well developed as older children (Thomas 1980, Joiner, Messer et al. 1998, Yan, Thomas et al. 2000), and younger children produce slower movements with more jitter and higher variability than compared to older children (Yan, Thomas et al. 2000). Differences in motor performance speed and accuracy have been observed between 5, 7 and 9 year olds (Kerr 1975). In another study comparing performance of arm movements of subjects aged 6,9, 24 and 73, (Yan, Thomas et al. 2000) found that children 6 year old produced slower and more inaccurate movements, similar to the 73 year old adults; this was in contrast with the 9 year olds which produced similar performance with 24 year olds.

Some movement tasks require coordination not just between muscle groups, but also between muscles actions and visual perception (Payne and Isaacs 2002). When throwing a ball or striking a ball with a bat, the activities involve a high degree of muscle coordination and real-time adjustments based on visual perception. These skills develop throughout childhood. Children 4-5 can catch a ball if thrown gently from a short distance, but at this age, children have trouble striking with long-handed equipment, such as bats and hockey sticks. At ages 6-7, children show improved catching and striking skills, being able to strike large objects or slowly-moving balls. By age 8, they can catch balls thrown at different parts of their body, and begin to show mature form in skills of throwing and catching; however, at this age children still have trouble with skills involving hitting moving objects with paddles, bats and sticks.

The above tasks are dependent on children's visual perception, which develops until late childhood. Children's ability to track moving objects with their eyes develops until age 9 (Gallahue and Ozmun 1998). Children's ability to distinguish objects from background stabilizes only by age 13 (Gallahue and Ozmun 1998), and acuity for static and moving objects is developing until roughly 12 years of age (Gallahue and Ozmun 1998). Further, young children have trouble estimating distances to objects, understanding trajectories of moving objects, and coordinating their limbs to intercept moving objects (Gallahue and Ozmun 1998, Payne and Isaacs 2002).

The development of hand dominance is also an important factor to consider in the context of handheld technology, because fine motor performance has been shown to be better for the dominant hand (Perry and Hourcade 2008). Handedness is only clearly established after 10 years old (Belmont and Birch 1963). Before this age, hand preference may not be visible in some children – for example, in a study comparing across ages, (Belmont and Birch 1963) found that hand dominance was visible in 85% of children age 8½ years, while only 70% of children 3-6½ years old had established dominance.

Gender differences do not appear significant for young children's motor performance. Children's physical development is similar for both genders during the years of early childhood, and studies of young children's motor performance do not detect differences between genders in young children (Kerr 1975, Vasta, Regan et al. 1980).

The development of motor skills through childhood hints at potential issues that may be encountered by 7-9 year old children as they use handheld AR interfaces. It is expected that children's precision will increase with age, and younger children will show slower and more imprecise movements, with higher variability in their movements. Gross motor skills are expected to be developed, and children should not have issues moving their whole body around the AR space. However, the developing ability to coordinate movements will cause an impact on some tasks, such as when children are moving while holding an AR object within view, tracking or touching virtual objects that are moving, or when children must hold the phone with one hand and perform another task with the other hand (such as touching the screen or manipulating a paddle). Younger children are expected to be worse at these tasks, leading to higher inaccuracy, and events such as dropping the device, or frequent loss of tracking while manipulating the device.

Age differences can be observed in children's interactions with various pointing devices, and studies show that pointing accuracy increases with age (Hourcade 2006). Compared to younger children, older children and adults are quicker, less jagged and more accurate at using input devices such as mice and touchscreens (Joiner, Messer et al. 1998, Oehl, Sutter et al. 2007, Anthony, Brown et al. 2012), and younger children learn slower (Scaife and Heckler 2010). These studies correlate with motor development literature that suggests children's motor skills continue to develop into early teens.

Although no studies report on children's pointing accuracy in with handheld augmented reality, it is expected that similar effects will be observed. Specifically, it's expected that younger children will be slower and more imprecise than older children in their ability interact with the screen. However, it is worth noting that in interacting with handheld augmented reality, children will perform movements with small muscle groups in their hands and fingers (pointing and holding the device), as well as with large muscle groups in their body (leaning and walking in a larger physical space), and inaccuracies in interaction may be compounded when children must perform both movement types at once (for example when interacting with a virtual object that is moving in physical space). Additionally, when children move the device while attempting to perform an action, they will cause virtual items to change position on the screen, thus hand-eye coordination will be required if children cannot hold the device steady.

2.2 INTERACTION TECHNIQUES

There is a long tradition of research comparing between different input mechanics. Early studies have focused on comparing user performance between input devices, such as mouse, joysticks, trackpads, etc. (Card, Mackinlay et al. 1990). Other studies have compared possible interactions with only one input device, such as comparisons of point-and-click vs. drag-and-drop mouse interactions (Inkpen 2001, Hourcade 2008). Interaction technique research has flourished in the domain of virtual reality, where researchers have created testbeds and taxonomies for systematically comparing interaction techniques (Bowman and Hodges 1999). In this section, I will discuss various aspects of designing interaction techniques.

2.2.1 Motor Complexity

Many studies have investigated children's performance on interaction techniques for PC based applications, and performance appears to decrease with an increase in the complexity of the interaction gesture. When children are required to hold down a mouse button while moving the device (such as required by a drag-and-drop interaction), they are slower, produce more errors, and indicate lower preference for the interaction than compared to dragging without holding a button (Berkovitz 1994, Strommen 1994, Inkpen 1997, Joiner, Messer et al. 1998, Inkpen 2001, Elliott, Hansen et al. 2004, Grossman, Hinckley et al. 2006, Oehl, Sutter et al. 2007, Hourcade 2008). These performance differences are likely caused because the hold-and-drag interaction requires coordination between multiple muscles, and demands endurance for holding a button pressed (Strommen 1994, Oehl, Sutter et al. 2007). While using touchscreen devices, children also have trouble with more complex gestures such as double-tapping, than compared to single touching (McKnight and Fitton 2010). These observations are expected to transfer to handheld AR usability, whereby movements that require more muscle coordination will lead to lower performance.

The manner of coupling between input and output is a factor that has been noted in studies investigating children's performance on various input devices. There are two aspects to consider regarding input-output coupling. First, is the difference between "input and output space": in traditional desktop interfaces, the input is performed on a user's desk, while the output is displayed on a screen. With touchscreen interfaces, these spaces have been combined, a factor which likely contributes to the ease of use of these interfaces. Second is the directness of the "mapping" between input movement with output movement: a mouse interface is a fairly direct mapping, whereby the path of the mouse movement on the table corresponds to the path of the mouse pointer on the screen; on the other hand, a joystick is a more indirect mapping, whereby nudging the joystick and holding it in place causes the mouse pointer to move on the screen and continue moving in a straight line until the joystick is released.

(Revelle and Strommen 1990) suggest that joysticks and trackballs are more difficult to learn than mice because of the indirect mapping between input and output of these devices. This may explain why touchpads are easier to use than joysticks (Oehl, Sutter et al. 2007), and mice are easier to learn than joysticks and trackballs (Revelle and Strommen 1990). The issue of indirect mapping is apparent in webcam applications where upward movements toward the screen lead to downward movements on the "magic mirror" screen, and this has been reported as problematic for children (Hornecker and Dünser 2009). Handheld augmented reality applications do not suffer significantly from issues related to motion mapping – there is a direct mapping between input and output motions; although the offset between the camera and user's eyes could present issues in some situations (Kruijff, Swan et al. 2010).

2.2.2 Touchscreens

Revelle and Reardon (Revelle and Reardon 2009) discuss general guidelines useful when designing touchscreen applications for young children. Among other guidelines, the authors warn that children may tilt device while playing games, and that children unintentionally touch the edge of the screen while holding handheld devices. Although these guidelines have been developed for traditional handheld games, they are transferrable to the domain of augmented reality.

Unintentional tilting of the device may cause issues in handheld augmented reality, since moving the device causes virtual items to move on the screen, and may cause the camera to lose tracking. For children who may not be able to hold their hands steady, this may cause inability to accurately point at the game items, resulting in confusion and frustration.

Issues related to unintended touch have been reported in other studies of children interacting with handheld devices. Children have been observed to touch the edges of the screen or to use multiple fingers while touching a screen location (Romeo, Edwards et al. 2003, McKnight and Fitton 2010, Anthony, Brown et al. 2012), and children may tap multiple times on an item even after the action has been registered by the application (Anthony, Brown et al. 2012).

Unsurprisingly, children's touch accuracy decreases with target size, and children find small icons difficult to touch and cause for frustration (Romeo, Edwards et al. 2003); these observations mirror studies of Fitts' Law on mouse-based interactions. Additionally, studies with adults have found that accuracy increases when targets are closer to the edge of the screen (Perry and Hourcade 2008); this effect is expected in children's interactions, but has not been tested.

It is expected that in handheld augmented reality applications, unintentional and inaccurate touch may be more severe than in non-AR handheld applications. If children need to constantly hold and move the device while aimed at a trackable target, their hands may unintentionally touch the edge of the screen more frequently.

2.2.3 Handedness

Handedness has been noted as an important consideration for handheld interface usability. When adults perform a touchscreen interaction with their dominant hand, their interactions are faster and more accurate (Hancock and Booth 2004, Inkpen, Dearman et al. 2006, Perry and Hourcade 2008). When observing various non-touchscreen input devices, (Kabbash, MacKenzie et al. 1993) observe that use of dominant hand in adults leads to improved performance in tasks requiring precision, such as when working with small targets and small distances. Interfaces may be specifically designed to appeal to the user's dominant hands, such as in the case of GUI scrollbars (Inkpen, Dearman et al. 2006). Unlike desktop GUI interfaces, handheld AR interfaces do not typically offer detriments for left-handed users. However, the choice of hand will affect user performance in handheld AR interactions, and it may be difficult to judge which hand will lead to best performance some interactions – for example when an interaction requires the use of both hands, such as holding a handheld phone with one hand and moving a paddle with the other hand.

2.2.4 Interaction Tasks

In the domain of virtual reality, (Bowman and Hodges 1999) have introduced a task-based taxonomy for classifying types of interactions that a user can perform in an application. There are three main classes of interactions, according to the task they achieve: Selection (se, indication of which virtual object should be the target of an action), Manipulation (e.g., positioning or orienting objects), and Travel (i.e., moving the user's point of view in the virtual space) (Bowman and Hodges 1999, Bowman, Johnson et al. 1999). The tasks of "selection" and "manipulation" are applicable to handheld augmented reality, while the task of "travel" is delegated to moving the handheld device. In my research, I am primarily interested in studying techniques for "selection", because selection is necessary for interaction with any virtual environment (Bowman, Kruijff et al. 2005).

An application can provide the user with many different methods for achieving these tasks. For instance, many interaction techniques have been developed for the task of "selection" in virtual environments (Poupyrev, Billinghurst et al. 1996, Pierce, Forsberg et al. 1997, Bowman, Johnson et al. 1999, Vanacken, Grossman et al. 2009). The choice of interaction technique depends on the application environment and task requirements (Bowman 2002). For instance raycasting techniques, which involve selecting objects by shooting a straight ray from the user's mobile touchscreen or VR glove, are optimal for tasks requiring speedy selection, but their performance suffers when virtual objects are closely clustered or occluded (Bowman 2002). Interaction techniques have been widely developed for head-mounted and webcam AR systems (Thomas and Piekarski 2002, White, Feng et al. 2009, Ha and Woo 2010, Oda and Feiner 2012).

In handheld augmented reality applications, the most common selection interaction techniques are crosshair-based selection (i.e., raycasting from the center of the screen), or finger touch-based selection (i.e., raycasting from a point indicated by the finger on the touchscreen). A variety of applications using these interactions exist for children (as discussed in Chapter 2.3.1); however, children's use of these interactions have not been systematically studied. Research on handheld AR interaction techniques has mostly focused on adults, using more complex interaction techniques for selecting in crowded virtual spaces (Olwal and Feiner 2003, Mossel, Venditti et al. 2013), manipulating virtual objects in large spaces by locking them to the screen (Guven, Feiner et al. 2006, Mossel, Venditti et al. 2013), selecting objects by touching with virtual objects locked to the screen (Tanikawa, Uzuka et al. 2015), or selecting and manipulating objects by using a real hand in front of the handheld camera (Hürst and Van Wezel 2013, Seo and Lee 2013, Lakatos, Blackshaw et al. 2014, Kim and Lee 2016). Only one research investigation has been found to compare the basic crosshair-based and finger-based selection techniques (Hürst and Van Wezel 2013). When testing adult users, the authors found crosshair-based interaction was slower, but it provided benefits in precision, and users found it to be more fun. When comparing children's interactions with crosshair-based selection, it is expected that young users will also show a slower task completion time. It is unclear if children's developing physical skills will allow them to show increased accuracy with the crosshair. It may also be that the crosshair interaction will be too hard for some children, thus showing decreased enjoyment levels. No existing studies have investigated the suitability of interaction techniques for children in immersive VR or AR spaces, for either handheld and non-handheld platforms. However, a small variety of handheld-AR interaction techniques are present in existing games, as will be discussed in section 2.3

2.2.5 Testbeds

Interaction technique research is typically done in simple software environments, as seen in typical Fitts' Law studies (Song, Clawson et al. 2011). These software environments for measuring user performance on repeated tasks have been called testbeds in (Bowman, Johnson et al. 1999), and I will use this terminology in this thesis. The research design is typically within-subjects, and involves exposing participants to repeated tasks which must be completed using specific interaction techniques (Bowman, Johnson et al. 1999). The experimental software logs quantitative user data such as speed and accuracy, and additional measures such as perceived ease of use or preference are collected using post-task surveys and interviews. In my research, I intend to replicate this study methodology, but I also wish to extend it by collecting qualitative data about usability issues encountered during the participant's interaction.

2.2.6 Effects of Practice

Learning effects are important considerations when observing child performance in various interaction styles. Observation of individuals repeatedly performing a variety of tasks has shown that learning follows a power law relationship (De Jong 1957, Newell and Rosenbloom 1981). This model has been shown to account for learning of input devices (Sutter 2007). This model indicates that user performance improves drastically in the early stages when users are learning an interaction, and is useful in comparing learning across modes of input.

The speed of learning depends on the user's age, expertise, and types of interaction device used. Novice performance on trackpoint and touchpad devices has been shown to achieve stabilization after 960 trials in (Sutter 2007), and research with other input devices indicates stabilization after 1400-1600 trials (Trankle and Deutschmann 1991). Adults and 5-year-old children learn to use touchscreens at the same rate, but learning rates are different for mouse interactions (Scaife and Bond 1991). It's likely that the difference in learning is because touchscreen interaction is more natural and familiar than manipulating a mouse. Due to the nature of augmented reality, some interactions are expected to be will be easier to learn since they are directly transferrable from the real world (for example, touching a virtual item on the screen is similar to pointing at an image on a piece of paper); however, other interactions will be more unfamiliar to users and may require more learning (for example, aiming at virtual items by panning the handheld screen).

The generalizability of results from short-term learning tasks is further constrained by observed differences between novices and experts (Joiner, Messer et al. 1998, Sutter 2007). In the study by (Sutter 2007), novices were observed to have slower performance than compared to experts, even after prolonged practice and visible stabilization of performance. This indicates a limitation with interaction-technique studies, in that observed stabilization of task performance in the short term does not necessarily indicate that users have reached expert performance.

2.3 CHILDREN AND AUGMENTED REALITY

In the following discussion, I will describe existing AR applications for children, for the contexts of handheld, webcam, and head-mounted-display (HMD) applications. I will present different interaction techniques available for each platform, and will discuss the usability issues experienced by children with each type of interaction.

2.3.1 Handheld Interactions

Handheld-AR applications require the user to hold a mobile device, such as smartphone or tablet, and look at a trackable 2D image or 3D object. Children's AR games have been popularized by devices such as the PlayStation Portable and the Nintendo 3DS. These devices are specifically designed for entertainment, and the user interaction with AR content is through the use of physical buttons. In recent years, however, the widespread adoption of camera-enabled smartphones has caused an increase in AR games designed for touchscreen. I am focusing my research on touchscreen-based AR applications, since the continuing popularity of smartphone devices will make AR accessible to large amount of children.

Handheld AR applications are split across two groups, depending on the AR technology used. One set of applications requires the user to be looking at a trackable marker. This design constrains the user's movements, requiring the device camera to be pointed at a specific trackable image. This design also allows for complex spatial movements, such as moving closer/farther from the game pieces, or changing one's orientation relative to the pieces. The other set of applications uses the mobile device compass and accelerometer, thus users can look up/down and rotate left/right to view virtual items in any direction of their environment; however, physical movement does not change the user's relationship to the items. The user has more freedom to look in a wider area, since they are not bound to looking at the marker, but the interactions require less spatial complexity, since users cannot move closer to the game objects.

No existing research has compared children's experienced usability issues between marker- vs. compass- based applications. However, these differences have an effect on user experience and usability issues: compass-based experiences create more inaccurate responses to user movements (since mappings between user's hand movement and the movement of virtual items are affected by accelerometer errors), and marker-based experiences are more prone to tracking loss (since the game cannot function if the device camera doesn't detect the image in its field of view). Some interaction mechanics are applicable to both types of handheld AR systems (e.g., the user can touch the screen to select virtual items, regardless of whether the system is using compass- or marker-based tracking), but some interactions can only be implemented currently in marker-based systems (e.g., in marker-based systems, the user can look around an object to touch it from the side; this is not possible in compass-based systems since the system cannot know if the user's spatial location has changed). I am interested in studying marker-based systems in my research because they allow exploration of a larger variety of interaction mechanics.

Existing handheld AR games for children make use of several classes of interaction mechanics. The wide majority of these mechanics have not been empirically evaluated, and it is unclear what usability issues are experienced with each interaction, or at what age children can use the interactions. Table 2.3.1. offers a summary of existing interaction techniques for the task of selecting virtual items. The applications listed in the table are available on the iTunes store, and/or have been presented in academic publications, and/or have been created by myself.

Table 2.3.1. Handheld-AR applications, organized by motor and spatial complexity of interaction. Applications marked with star (*) are marker-based, while others are compass-based.

SPACE / MOTOR	**Don't need to understand/move in 3D space.**	**Need to understand/move in depth, without changing orientation**	**Need to understand/move in whole space, changing orientation**
Hands move together, no screen touch	Look at an object to trigger action [Band Aid*, (Mott, Bucolo et al. 2008)*]	Screen proximity [Mountain Rescue*]	Virtual attached object [Nerdherder*] Screen proximity +orientation [Bacteria Snap 2*]
Touch screen edge, aiming with crosshair.	Crosshair aim, touch anywhere [Real Strike, Zombies Everywhere, Aliens Everywhere, Shutgun 3D, Rocks in Socks*, Transformers3, Real Shooter, Paintball Arena, Alien Attack, Ghostbusters, Fly Hunter, GyroShootAR] Crosshair aim, touch edge button [AR Earth Invasion, AR Defender*]		Crosshair aim, touch anywhere: [ARC Popper*]
Touch screen interior.	Touch specific item on screen: [Wikitude, Zombies Everywhere, Anomaly Ultimate, War of Worlds, Mushabellies*, TapCloud, Chromium Wars*, Spintopia*, Morrison*]	Touch location on plane [Puppy Dog Fingers, Bunny Fingers, War of Worlds Pocket Jets, Chromium Wars* AR Defender 2*] Touch specific item and drag to location [Mushabellies*]	
Hold screen with one hand, move other hand into space in front of camera		Reach with marker in space, not very precise [Puppy Plus*]	Reach with paddle, precise [Bacteria Snap 1*]

The interactions are organized according to two dimensions: Motor Complexity and Spatial Complexity. The Motor Complexity axis indicates the complexity of hand movement, taking into account precision and coupling between hand movements. In the most easy to manipulate games, the user can hold the device with both hands without being required to touch the screen (e.g., in Nerdherder (Xu, Mendenhall et al. 2012), the users control the game through the movement of a virtual doughnut which is attached to their device). A more popular interaction technique is using a crosshair, where the user can hold the device with both hands but must touch somewhere on the screen in order to indicate that a targeted object should be acted upon (e.g., in GyroShootAR, users orient their device toward floating aliens, then touch the screen to shoot a laser gun). Another popular, yet potentially more difficult, interaction technique is directly touching virtual items on the screen; this interaction requires the user to hold the device with one hand, while the other hand performs the touch gesture (e.g., in TapCloud, users must touch on virtual clouds in order to pop them). Finally, a more complex interaction technique requires the user to hold the device with one hand, while using the other hand to reach into the virtual space (e.g., in Bacteria Snap 1 (Radu and Bujak 2012), users must neutralize a bacterium by touching it with a virtual antibody attached to a physical paddle).

The Spatial Complexity axis organizes interactions according to how deeply a user must understand and potentially navigate space in order to interact with the game. In most existing games, the user does not need to change their physical location, and is not required to understand spatial relationships between the elements of the game (for example, in Zombies Everywhere, the user must tap on the screen location of zombies in order to destroy them; whether zombies are in front or behind each other is not important, and the user could play the game without understanding the relationship between the zombies and the user's location). Other games require users to understand and move in relation to depth, but do not require users to change their orientation relative to game objects in order to play (for example, in Mountain Rescue (Radu and Bujak 2012), users interact with virtual items by moving the device close to the items they wish to act upon). Finally, some games may require users to change their physical location as well as orientation, requiring an understanding of the relative positions of game objects (for example in Bacteria Snap 2 (Radu and Bujak 2012), before children can neutralize a bacteria, they must first move close to the bacteria, then turn their device to match the orientation of the bacteria).

The majority of these interactions have not been empirically studied, but scattered findings hint at the possibility of usability issues increasing in severity as complexity increases among the motor axis. The study by (Mott, Bucolo et al. 2008) and my own observations of Mountain Rescue (Radu and Bujak 2012), indicate that young people can easily interact with a system that requires holding the device with both hands. When requiring the users to touch the screen, some issues appear. The study by (Morrison, Oulasvirta et al. 2009), reporting on observations from users aged 7-50 using a system where users must touch virtual items on screen, indicates that users sometimes drop the device while interacting. My own pilot observations of children 5-6 playing with Spintopia (Radu and Hewner 2011), found that young users occasionally drop the device, and are inaccurate in their pointing gestures. Furthermore, when users are required to perform independent actions with both hands, young users appear unable to perform these interactions. In the pilot of Puppy Plus (Radu, Hanlon et al. 2011), children were observed to put down the device, then perform physical manipulation of game pieces, then pick up the device again. Finally, while playing with Bacteria Snap (Radu and Bujak 2012), children 5-7 were generally unable to perform the interaction without assistance.

Increasing the spatial complexity of the interaction appears to lead to increase in usability issues. In observations of the Mountain Rescue (Radu and Bujak 2012) game, children around the age of 6-8 were observed to be able to move the device closer/far from the marker without severe issues. On the other hand, in a pilot observation of the ARC Popper game, which requires children to aim at virtual items located inside virtual containers, children 7 and 9 were observed to have trouble moving their bodies to look inside the virtual boxes. Additionally, in Bacteria Snap 1 (Radu and Bujak 2012), a game where were required to match the spatial location and orientation of a physical paddle with a game item, children 5-7 were unable to perform this interaction. These indicate that challenges may be encountered when children must move their hands or whole body in relation to virtual game elements. In the limited literature about children's interaction with handheld augmented reality, an issue is reported regarding movement in large spaces. (Dunleavy, Dede et al. 2009, Morrison, Oulasvirta et al. 2009) observe that when participants play geo-located outdoor games, participants exhibit a high focus of attention on the game while moving in the physical world, leading to potential safety issues such as bumping into objects or walking into traffic. In my research experiments, I expect this high-attention effect will be present as children move around, and I intend to investigate other issues that arise from the spatial complexity of the interaction.

Interaction techniques for handheld AR have not been systematically compared. In my studies, I will select several interaction techniques, varying among motor and spatial complexity axes, and I will investigate how children of different ages react to these techniques.

2.3.2 Webcam and Head-Mounted Display Interactions

In this section, I will describe the different kinds of techniques used for interacting with webcam and head-mounted display (HMD) AR systems, and the usability issues encountered around these approaches. I present the interaction techniques starting with the simplest, and moving to increasing complexity.

The majority of existing publications on children's use of AR are in the domain of webcams, and several other publications discuss children's AR applications for head-mounted displays. The webcam setup captures the user's actions in a physical space, and displays the augmentation on a computer display; the interaction space is typically a flat surface such as a table with a webcam placed above it, while the AR display is a monitor or a projection screen. In head-mounted display applications, children wear a helmet equipped with an internal display and an external camera, and look at physical objects which they manipulate in order to influence the augmentation.

The most frequently used interactions in webcam AR applications are based on using physical props, such as paddles or cubes that contain detectable AR markers. The simplest kind of interaction involves showing the marker to the application. Some applications use this interaction to permit unconstrained play (Kern, Stringer et al. 2006, Juan, Llop et al. 2010), other applications rely on this interaction as a way for students to select answers to questions (Freitas and Campos 2008, Chien-Yu, Chao et al. 2010), and a variation on this technique is the hiding of markers to create a game action, such as playing a sound (Correa, de Assis et al. 2007). These kinds of interactions are simple to do because they require children to generally move or hide markers within the field of view of the camera, and most studies do not report usability issues. One exception is, (Correa, de Assis et al. 2007), in a system intended for physical therapy for children with physical palsy, has found that repeated hand movements became tiring after about 10 minutes for the participant child (unknown age).

A slightly more challenging interaction technique is requiring the user to move a physical object to a specific location in the physical space, such as putting an AR marker close to another marker in order to trigger a game action. It is worth noting that this interaction does not require the user to look at the AR display while performing the action – instead, the user can focus all his or her attention on the physical space. It is implemented in a variety of webcam applications (Richard, Billaudeau et al. 2007, Zhou, Cheok et al. 2008, Marco, Baldassarri et al. 2010, Campos and Pessanha 2011). These kinds of physical actions are not reported to be problematic for children, although children 3-5 years old were observed to exhibit imprecise movements (Marco, Baldassarri et al. 2010).

We begin to see more usability issues appear when children are required to perform physical actions at the same time as they look at the PC monitor displaying the augmentation. In the study of WizQubes (Zhou, Cheok et al. 2008), children were required to rotate physical cubes in order to change the current game level or currently selected game item; in the study observations, some children did not watch the screen while performing this action – this may indicate a temporary curiosity about the physical cubes, or it may be due to a limited ability to divide attention between the physical and virtual spaces. Another observation from (Zhou, Cheok et al. 2008) was that children covered the AR markers, even after being instructed about how the vision-based AR technology works – this may indicate limitations in attention or memory, or that children have accurate models of how the technology operates. Several applications require children to perform more complex actions while watching the screen, such as moving physical markers or paddles towards items that are only visible in the AR view (Theng, Mei-Ling et al. 2007, Hornecker and Dünser 2009, Radu and MacIntyre 2009, Marco, Baldassarri et al. 2010). For example, in the biology education application from (Theng, Mei-Ling et al. 2007), one of the activities involves watering a virtual plant by moving a physical paddle whose AR representation is a virtual water sprinkler. The study reported that children had difficulties operating the paddle, though the specific nature of the difficulties is unclear. The study of the interactive storytelling system from (Hornecker and Dünser 2007, Hornecker and Dünser 2009) indicates that children experienced confusion with spatial directions – potentially because upward hand movements resulted in downward movements on the screen. In the same studies, it was found that children experienced attention issues (children were observed to ignore verbal instructions) and memory issues (children expected that the system would respond to real-world interactions, such as banging items together to break them, even after being trained otherwise) (Hornecker and Dünser 2007, Hornecker and Dünser 2009). It is possible that requiring children to divide attention between physical and virtual spaces while manipulating physical objects, placed high cognitive load and caused focused attention toward the game experience, causing children to ignore audio stimuli and previous instructions. In handheld-AR the input and output spaces are integrated, unlike in webcam systems; however, in handheld-AR children are observed to experience issues with highly focused attention (Dunleavy, Dede et al. 2009) and difficulties of physically manipulation (Radu and MacIntyre 2012), and it is expected that memory issues with expectations of technology will also be experienced.

Interaction with two hands at the same time is required by some webcam-based systems, and by all HMD-based systems for children. The system in (Nischelwitzer, Lenz et al. 2007) requires users to push buttons on a book in order to affect the AR experience; it does not report usability issues with this interaction. In the webcam WizQubes system (Zhou, Cheok et al. 2008), children must bring two large physical cubes close to each other in order to activate a game action; this interaction is simple to perform as it requires gross motor skills which develop early in childhood, but usability issues were experienced as children accidentally covered the cubes causing loss of AR tracking. The webcam system presented in (Andersen, Kristensen et al. 2004) also requires 2-handed interactions, as children must physically connect two LEGO pieces in order to activate game battle sequences; children exposed to these systems were able to perform these actions, but some reported that the actions took too long to perform and disturbed the game flow. In the HMD system presented in (Juan, Toffetti et al. 2010) children must move cubes close to each other to trigger animations; although no specific usability issues were reported with this application, children reported the AR condition less easy to use than the non-AR condition. In (Juan, Beatrice et al. 2008), children had to open and close physical zippers to trigger test answers; the study observed that children had trouble manipulating the zipper interaction, and this may potentially be due to children's developing fine motor skills and bimanual coordination skills. A more complicated HMD-based 2-hand interaction is presented in, (Zhou, Cheok et al. 2004), where children must fold and unfold 3D cube puzzles to change the levels of a game; no evaluation was performed on this application so it is unclear what kinds of usability issues children experience with this interaction. The results in these systems indicate that children experience interaction issues when being required to perform movements with both hands independently; I expect these issues to transfer to handheld-AR interactions where independent hand movement is required, such as users needing to hold phone with one hand while manipulating a paddle with the other hand.

Other webcam-AR applications make use of general body movements as interaction techniques. These are typically forward-facing cameras, and the user sees their own reflection in an augmented magic mirror. For example, in (Sony Computer Entertainment), users wave their hand to pet a virtual character or to bump virtual items, and in (Good, Romero et al. 2008) the user's joints are connected to a virtual character's joints. No usability issues are reported with these interactions; for handheld-AR it is expected that children will be able to move their body to reposition the device in a general location, but issues may be observed if the interaction requires precise aim.

These observations indicate that children experience a variety of usability issues, and their severity increases with the complexity of the interaction. In (Radu and MacIntyre 2012), I summarize the usability issues encountered by children, along with developmental psychology concepts which might account for the issues.

CHAPTER 3

PILOT STUDIES OF EDUCATIONAL PROTOTYPES, AND INITIAL USABILITY FRAMEWORK

3.1 OVERVIEW

Three components of my previous work have influenced the creation of the main research experiment described in Chapter 4. First, I have built and informally evaluated several educational AR applications for children, each requiring the use of a specific interaction technique. These informal investigations have informed me of the suitability of different interaction techniques for children of various ages, and have indicated different types of usability issues that children may encounter in handheld AR games (see Chapter 3.2, and (Radu and MacIntyre 2009), and (Radu, Xu et al. 2013)).

Second, I have constructed a usability framework for designing age-appropriate AR experiences for children. The framework describes usability issues encountered by children (as observed in existing AR literature and in my own informal observations), presents developmental psychology constructs that explain the usability issues, and predict potential new issues (see Chapter 3.3, and (Radu and MacIntyre 2012)).

Finally, I have analyzed the educational benefits of augmented reality. Through several theoretical papers, my colleagues and I have discussed how student cognition can be influenced through the affordances of this medium, presenting opportunities for learning that are different than other media (see Chapter 3.4, and (Radu 2012, Bujak, Radu et al. 2013, Radu 2014, Radu, Doherty et al. 2015, Radu, McCarthy et al. 2016)

3.2 EXPLORATION OF INTERACTION MECHANICS

I have constructed six different applications for children, each using its own interaction mechanics, in order to explore the design space of educational AR applications for young children. Five applications are handheld AR games: Spintopia (an art game requiring the user to touch and drag items on the screen) (Radu and Hewner 2011), Puppy Plus (a math game requiring children to hold the phone while manipulating physical wooden pieces) (Radu, Hanlon et al. 2011), Mountain Rescue (a biology game requiring users to hold a device with both hands while moving close to different virtual items) (Radu and Bujak 2012), Bacteria Snap (a biology game requiring users to hold the device with one hand while manipulating a physical paddle) (Radu and Bujak 2012), and ARC Popper (the experimental game used in my main study described in Chapter 4). The remaining application is a PC webcam application AR SPOT (an authoring environment based on MIT's Scratch, requiring users to use physical playing cards and other tangible props). Table 3.2.1. presents a summary of the interaction techniques in these applications. ARC Popper is described in Chapter 4, and the remaining applications and their associated pilots are described in the sections below.

Table 3.2.1. Interactions used in my handheld AR games.

SPACE / **MOTOR**	**Play by standing still (don't need to understand game as 3D space)**	**Need to understand/move in depth, without changing orientation**	**Need to understand/move in whole space, changing orientation while walking around gameboard**
Hands move together, no screen touch		Mountain Rescue (move close to virtual item to interact with it)	
Touch screen edge, aiming with crosshair.	ARC Popper level 3 (select by aiming the crosshair at fixed item, touch edge to select)		ARC Popper level 4 (interact by touching screen edge, need to move body around game)
Touch screen interior.	Spintopia (touch screen to select & move virtual items) ARC Popper level 1 (touch screen to select)		ARC Popper level 2 (touch screen to select, need to move body around game)
Hold screen with one hand, move other hand into space in front of camera		Puppy Plus (place physical objects on marker to trigger virtual interaction)	Bacteria Snap (position and orient a physical paddle close to a virtual object)

Informal observations were associated with all of these applications. The observations involved participant ages between 3 to 13 years old. I observed children's ability to use each of these applications, and gathered usability issues associated with the application. The findings are summarized in Table 3.2.2., and described in the sections below.

Table 3.2.2. Problematic behaviors exhibited by children interacting with my handheld AR games.

Investigation	**Interaction Type**	**Findings by Participant Age**
AR Scratch	Webcam, mini-paddles	9-11 yo: Can generally use Trouble intercepting moving objects.
Spintopia	Handheld, touch/drag	3 yo: Cannot hold device, frequent drop. 6-7 yo: Can hold device Imprecise and jittery finger precision. Highly focused attention to screen.
Puppy Plus	Handheld, wooden pieces	6 yo: Cannot use with two hands, frequently puts phone down Sometimes did not notice tracking loss. Hands became tired. 8 yo: Became quickly bored of game.
Mountain Rescue	Handheld, closeness	5-10 yo: Game usable by all ages. Potential spatial/attention issues.
Bacteria Snap	Handheld, rectangular paddle	5-10 yo: Younger children could not use the interaction. Older children (8-9 yo) could hold phone with one hand, while using paddle.

The interaction in Mountain Rescue, which children could perform holding both hands on the device, was usable by children as young as 5; the SPOT webcam interaction was usable by the participants age 9-11; the touchscreen-based interaction in Spintopia appeared usable at 6 years but indicated potential accuracy issues; the two-handed interactions in Puppy Plus and Bacteria Snap were difficult to perform at 6 years but potentially usable by age 9. These informal observations hint that children may be able to perform some interactions but not others, and their capabilities is dependent on age. As a result of these observations, I became interested in systematically investigating handheld AR interactions to determine at what age children can perform specific interactions.

From the various informal observations above, it was observed that handheld-AR interactions present a variety of challenges for young children. Table 3.2.3 lists these issues. The usability issues have been classified under children's developmental skill that potentially is the strongest influence on each issue. In the following sections, I present details about my investigations with these specific applications.

Table 3.2.3. Classes of usability issues experienced by children.

Usability Topic	**Investigation**	**Findings**
Physical manipulation	AR Scratch, Spintopia, Puppy/Monster Plus, MountainRescue, BacteriaSnap	(9-11 yo) Issues with intercepting moving items [webcam paddle] (6-7 yo) Imprecise movements [handheld] (6-7 yo) Difficulty with drag & drop using finger on touchscreen [handheld] (6 yo) Difficulty with 2-handed interaction, put phone down and drop [handheld] (9 yo) Easier to use paddle interaction (6-10 yo) Closeness interactions were easy for everyone
Understanding of space	AR Scratch, Puppy/Monster Plus, MountainRescue	Understanding of 3D spaces not matured until late (9-10 yo) Ok projecting 3D space into 2D monitor (6-10 yo) Difficulties looking at large space through small screen [handheld] (6 yo) Difficulty judging distance for closeness interaction (6 yo) Focused on screen rather than space
Understanding of interaction	AR Scratch, BacteriaSnap	Children like interaction metaphors to be realistic, familiar (9-10 yo) Expected interaction metaphors to be concrete/physical (6 yo) Do not seem to notice when tracking lost, not know how to repair it (9 yo-older) Seemed to more easily able to repair tracking in paddle interaction, than compared to younger kids
Age differences	AR Scratch, PuppyPlus, Spintopia, MountainRescue BacteriaSnap	Older children more able to recover from tracking loss. Older children more able to understand and work with Paddle. Younger children issues with Paddle, Pieces, and with Drag&Drop. All children able to use Closeness interactions.

3.2.1 AR SPOT: Augmented Reality Authoring for Children

AR SPOT (Radu and MacIntyre 2009) is a webcam-based authoring environment based on the popular Scratch programming environment for children. In AR SPOT (Figure 3.2.1), children can create their own AR interactive experiences. Children use a visual drag-and-drop programming paradigm, to program the actions of 2-dimensional characters that live on the video feed captured from a webcam. The characters can be programmed to respond to movement of physical AR game cards, through simple commands as "stick to the blue card", or "perform an action if the blue card touches the green card". Once the characters are programmed, the user can run the game, watching the action on the computer screen, and controlling it through physical interaction with the game cards.

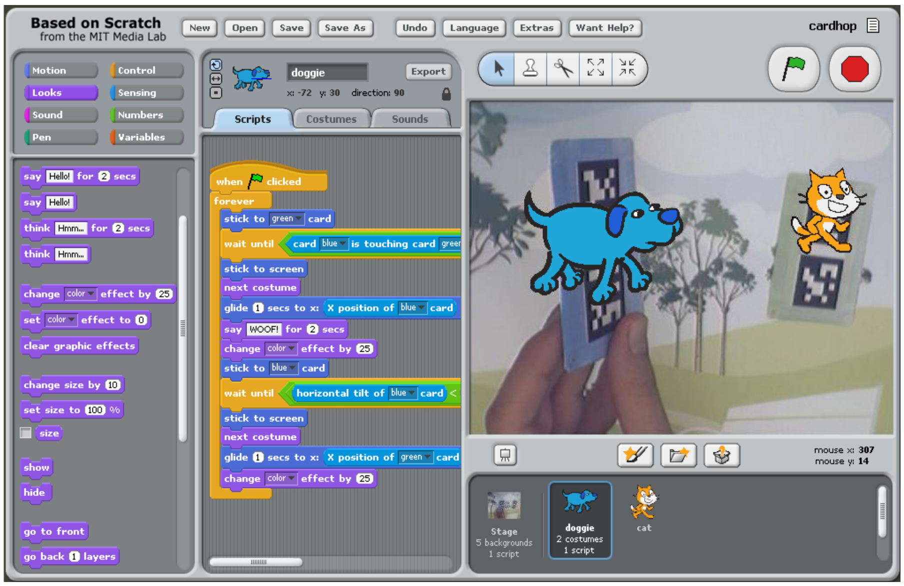


Figure 3.2.1. The webcam-based augmented reality authoring environment AR SPOT.

3.2.1.1 *Lessons Learned*

Although AR SPOT is a webcam-based system, its development and subsequent informal observations have provided insights that are applicable to the topic of children's interactions with handheld-AR systems.

The system design was informed by developmental psychology literature, which indicated that children's spatial cognition undergoes development in early elementary school years. Children as young as 5 years old are reportedly biased towards thinking in ego-centric perspectives, and, until about age 12, children have trouble thinking about 3D spatial relationships from the perspective of individual objects (Majid, Bowerman et al. 2004). Even though the AR technology would allow children to program interactions in three dimensions, the AR SPOT system was designed to operate in 2D screen coordinates in order to be accessible to wide audience of younger children. This knowledge further indicates that when children interact with 3D spaces through handheld-AR applications, they may have trouble understanding relationships between the objects, potentially leading to incorrect estimation of distances, incomplete memory for virtual object locations, or incorrect understanding of another player's perspectives.

Two studies were performed on the AR SPOT system. The first study was a pilot with two children aged 9 and 11. Among other findings (reported in (Radu and MacIntyre 2009)), we observed that children experienced some trouble interacting with the experience by manipulating the physical cards – specifically, it was observed that children had trouble using the cards to intercept virtual objects that were in motion. This skill relates to children's still-developing skills of dynamic visual acuity and interception of moving objects (see Chapter 2), and may indicate potential issues with interception of moving objects while using handheld-AR systems.

The second study involved a classroom of 12 students. The study was designed to identify what interactions children would expect from the AR SPOT system. Children did not interact with the system, but instead observed some use-case examples, and were asked to design games that they would wish to build using the system. Children used craft materials to design the games, and presented their games at the end of the study session. The games were coded in terms of the type of interaction technique that children desired in the game, and these interactions were presented in (Radu, Xu et al. 2013). Children favored games that were highly concrete (e.g., involved realistic characters performing actions governed by physical laws, as opposed to more abstract Tetris-style games) and interaction metaphors that frequently replicated the physical world (e.g., when a card was shaken or tilted, the character carried on it would frequently fall off). This bias towards physical realism in children's mental models of AR interaction has also been noted in (Hornecker and Dünser 2009), who indicated that even after children learned the repertoire of system interactions, they still expected interactions to have physical effects.

3.2.2 Spintopia

Spintopia (Radu and Hewner 2011) is a children's handheld-AR game for creating interactive art (Figure 3.2.2). This game requires the user to point their handheld device at a printed marker, and the primary interaction involves the user touching virtual items on the screen, or dragging virtual screen items. At the beginning of the game, the user creates a track by dragging their finger on the screen, resulting in a 3D train track on the marker. Once the path is created, a virtual particle shoots on the track, creating artistic effects. The user can create virtual effects at different points on the train track by touching screen buttons. Once the particle passes through the effect objects, fantastic visual effects are created in the AR space.

Figure 3.2.2. The handheld-AR art game Spintopia.

3.2.2.1 Lessons Learned

In a pilot study, three children aged 3, 6 and 7 years old interacted with the game. All children were engaged by the game's visual effects, but usability issues were observed relating to children's ability to manipulate and to comprehend the gameplay. The 3 year-old child was unable to hold the device appropriately, and frequently dropped it; furthermore, it was unclear if the child understood what was happening in the game. The 6 and 7 year olds could hold the device stable in order to observe the gameplay. However, the children soon treated the game as a drawing game – dragging their fingers on the screen to create random shapes, even if the AR tracking was lost. While the finger-drawing functionality was fun, the rest of the gameplay appeared too complex and/or uninteresting to these children; furthermore, children seemed oblivious of the fact that the AR tracking was lost and the game was not anchored to the real world anymore. These observations indicate that young children's attention can become highly focused on the on-screen effects from a handheld-AR game, to the point where children forget the hybrid AR space between the phone, as well as the connection between the game and the real world. Another observation from the study was that the children's finger movements on the touchscreen were jagged and imprecise, indicating that children would have trouble if asked to perform precise interactions in handheld-AR games.

3.2.3 Puppy Plus

Puppy Plus (Radu, Hanlon et al. 2011) was a handheld-AR game designed as an educational experiment for teaching addition skills to early elementary school children (Figure 3.2.3.). A puppy appears to be lost on an island in the middle of the ocean, and it is hungry for a specific number of doggy treats. The player's job is to keep the puppy well fed, by supplying it with the appropriate number of treats. The virtual island is anchored to a physical piece of paper, and the virtual treats are anchored to physical numbered pieces. The main player interaction involves watching the game world while moving the physical boat pieces. In order to feed the puppy, the player must pick up a physical boat piece and place it in a virtual dock; the level ends successfully when the numbers on the physical pieces add up to the number of doggy treats required by the puppy. This game was designed for two purposes: investigating the educational potential of application of handheld augmented reality, and investigating children's reactions to using tangible objects for interacting with a handheld-AR game.

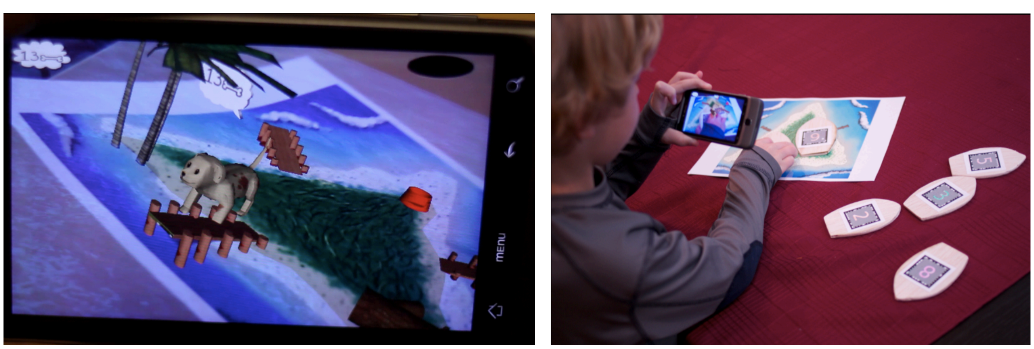


Figure 3.2.3. The handheld-AR game Puppy Plus, where the player must feed a puppy stuck on an island (device screenshot shown on left), by using physical numbered game pieces (shown on right).

3.2.3.1 Lessons Learned

I ran a pilot study of Puppy Plus with 6 and 8 years old children. The 8-year-old child became quickly bored of the game within a 1-2 minutes of gameplay, and did not interact with it for long, indicating motivation related issues that will be discussed shortly. The 6-year-old child interacted for an extended period of time (about 10 minutes), and his interactions revealed several items of interest.

The first finding was related to handheld-AR interaction techniques. During the design of the game, it was expected that children would perform two-handed interactions with the game: using one hand to hold the device aimed at the AR scene, and using the other hand to move the physical pieces. However, during the pilot study, the child was frequently observed holding the device with both hands; furthermore, in order to move the pieces, the child would put down the device, then move the physical pieces, then pick up the device and holding with both hands again. This indicated that children potentially have trouble performing handheld-AR interactions that require independent movement with each hand. This skill of bi-manual coordination is developing for young children (Payne and Isaacs 2002). Furthermore, the child reported that his hands became tired, indicating that children may experience muscle fatigue after 5-10 minutes of interacting with the handheld device.

Other findings included the fact that the child sometimes did not notice when the AR tracking was lost. Similar to the Spintopia pilot study, this indicates that children's attention may become very focused on the gameplay on the screen, to the point where they ignore the hybrid AR space between the device and the physical playing surface. Finally, motivational issues were experienced, as both children did not seem highly motivated to play the game. This is potentially because the low graphics quality of the virtual puppy, indicating that children may expect high realism when dealing with 3D games. Also, another motivational detriment may be the fact that the puppy would move around the virtual space, thus it was not always in full view of the small handheld screen; this potentially indicates that handheld-AR games will need to direct children's attention to important locations in the game space.

3.2.4 Mountain Rescue and Bacteria Snap

Mountain Rescue and Bacteria Snap (Radu and Bujak 2012) are two educational games designed to test different interaction techniques for handheld-AR games. Mountain Rescue (Figure 3.2.4 right) teaches children about animals that live at different altitudes on a mountain. Animals and their habitats appear overlaid on a vertical mountain poster. The player interacts with the game by moving their device close to virtual objects in the game: animals are picked up by when the player moves their device close to a rescue shelter, and animals are dropped off when the player moves the device close to an animal habitat. This interaction allows the user to hold the device with both hands, while they move close / far from the poster. Bacteria Snap (Figure 3.2.4 left) teaches children about antibodies and bacteria. In the game, various virtual bacteria appear on top of a physical poster; the player must neutralize them using an antibody, which appears on top of the player's physical paddle. To progress in the game, the player must neutralize each bacteria by bringing the antibody close to it, and rotate the antibody to match the orientation of the bacteria. This interaction requires the user to hold the device with one hand, and manipulate the physical paddle with the other hand.

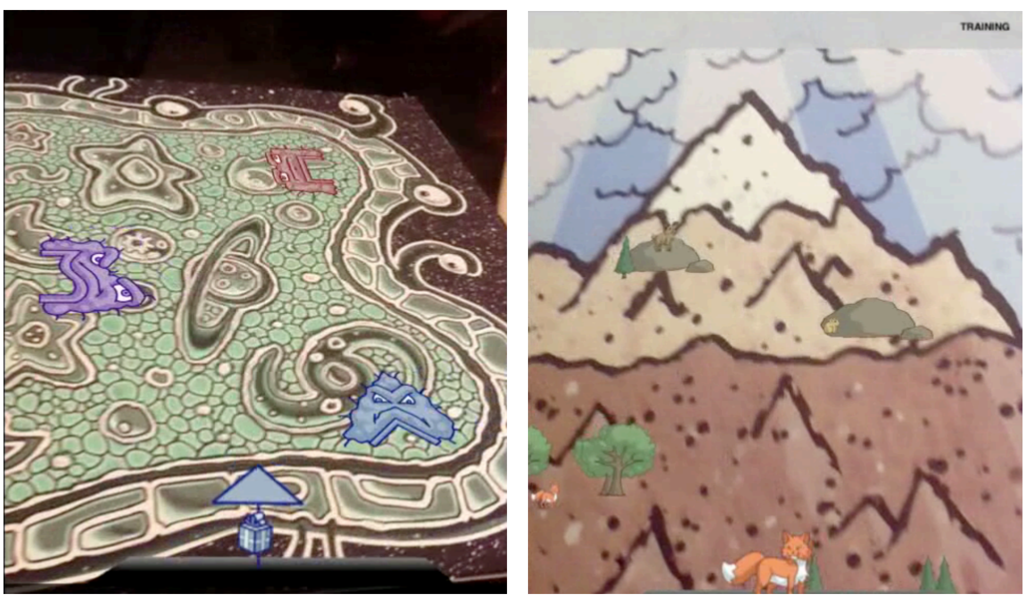


Figure 3.2.4. Bacteria Snap involves matching bacteria with antibodies (left). Mountain Rescue involves matching animals with their habitats (right).

3.2.4.1 Lessons Learned

Both these games were shown at the 2012 USA Science & Engineering Festival. Children of all ages visited the AR games booth, providing information about how different age groups react to the different types of game interactions. One major observation was regarding the ease of use of the two interactions. Most children were able to easily use the Mountain Rescue interaction of moving close / far in order to interact with animals and their habitats. In contrast, the paddle-based interaction of Bacteria Snap was challenging for children around the age of 5-6, but noticeably less challenging for children around the age of 8-9. Specifically, older children were able to hold the phone and paddle in a more stable manner, they were better able to maneuver the paddle to its required position, and they were better able to repair the AR tracking if the game lost tracking. These observations indicate that older children may have better motor skills, and better mental model of how the technology operates. One item of interest in the Mountain Rescue pilot was that, due to a software bug, the virtual game elements would change their position while the user was not looking at the game board, thus they would appear to be associated with slightly different physical locations; children did not appear to notice this, indicating highly-focused attention and/or inability to accurately remember spatial relationships of the game elements.

3.3 DEVELOPMENTAL PSYCHOLOGY AND AR USABILITY

In (Radu and MacIntyre 2012), I presented a framework for understanding AR usability for children. The paper presents a wide range of usability issues encountered by children, as reported by existing AR literature and my own informal observations. The issues are organized into categories based on developmental psychology factors that can explain the issues. The developmental psychology factors are from the areas of motor skills (bimanual coordination, hand-eye coordination, fine motor skills, gross motor skills, endurance), spatial cognition (spatial memory, spatial perception, spatial visualization), attention (selective attention, divided attention, executive control), logical thinking (abstract vs. concrete thinking) and memory (capacity and operations). These skills are illustrated in Table 3.3.1., along with example AR tasks that place high demands on the skills.

These developmental factors are all undergoing development in young children, thus have the potential to influence how children react to AR applications. Further, these factors might be useful to identify differences between children's AR usability at different ages. By understanding the developmental factors and their potential connection to usability issues, AR designers can use this knowledge to explain or predict user reactions to existing applications, and to design applications that are either usable or deliberately challenging.

Table 3.3.1. Developmental skills and potentially challenging AR designs.

Skill Type	Challenging AR interaction
MOTOR ABILITIES	
Multiple hand coordination	Holding a phone in one hand, and using other hand to move marker
Hand-eye coordination	Using a marker to intercept a moving virtual object
Fine motor skills	Moving a marker on a specified path
Gross motor skills & Endurance	Turning body around to look at virtual panorama Standing bent on a table, while playing handheld tabletop games
SPATIAL ABILITIES	
Spatial memory	Remember the configuration of a large virtual space, while using a handheld screen to see a limited view
Spatial perception	Understanding when a virtual item is on top of a physical item
Spatial visualization	Predict what virtual objects are visible by other people or virtual characters
ATTENTION ABILITIES	
Divided attention	Playing an AR game, and making sure to keep marker in view so tracking is not lost
Selective & executive attention	Playing an AR game while moving outdoors
LOGIC & MEMORY	
Remembering & reversing	Remembering how to recover from tracking loss
Abstract over concrete thinking	Understand that virtual objects are computer generated, and they do not need to obey physical laws

A limitation of this framework is that it does not provide any data about the specific AR designs which give rise to these usability issues. By presenting a broad set of guidelines supported by case studies and psychology theory, the framework has provided a foundation for constructing hypothesis-driven studies which can provide empirical data about relationships between usability issues and AR designs. In the study presented in Chapter 4, I employ both qualitative and quantitative methods to understand how different interaction technique designs influence children's performance and usability, and how developmental skills mediate these effects.

A second limitation of the framework is that it does not discuss how children's age influences usability with different AR designs. Differentiating usability issues by age was not possible due to the limited existing literature. However, understanding the relationship between age and usability is highly useful to AR designers, since children's products are typically designed for narrow age segments. In the research study detailed in Chapter 4, I investigate how performance and usability issues change across the 5-10 year old age range.

A final extension to this work is the generation of a coding scheme. The framework can be converted into a coding scheme for classifying child behaviors that are associated with usability issues (Table 3.3.2.). The coding scheme provides an association between observed child behavior and the developmental skill that potentially underlie that behavior. This scheme can be used as a basis for qualitative coding of video observations of children's use of different AR applications, providing a point of comparison between children's behavior between different AR designs or between different age groups. In the research study detailed in Chapter 4, this scheme has been used as a starting point for the qualitative analysis.

Table 3.3.2. Coding scheme for problematic child behaviors.

Usability Problem	
Short Description	**Observed Behavior**
Muscle Strain	User verbally complains about arm strain
Dropping	User drops phone
Shaking	User is shaking camera while looking at virtual item User is shaking camera while trying to touch User is shaking camera while looking and moving
Problem moving in space	While looking at game space, user has trouble moving closer to location / matching orientation of virtual item
Problem perceiving virtual positions in physical space	While looking at game space, user verbally can't indicate where virtual items are in physical space
Problem remembering virtual positions in physical space	Once user looks away from game, user doesn't know how to move/point camera at an off-screen virtual item
Problem perceiving game space	User verbally can't indicate which virtual item is in front of another virtual item
Can't describe point of view	User verbally can't describe another person's perspective of the game
Tracking loss away from marker	User loses tracking while playing because camera is pointed away from paper
Tracking loss too close to marker	User loses tracking while playing because camera is too close to paper
Tracking loss too fast	User loses tracking while playing because device moves too fast
Problem recovering tracking	User has trouble recovering from tracking loss (e.g., does not go to "optimal distance")
Problem remembering interaction mechanic	User does not appear to understand interaction mechanic (e.g., while aiming with crosshair, taps on virtual items instead of aiming toward them)
Ignores instructions	User ignores instructions from the game or researcher or parent.
Ignores physical environment	User bumps into physical objects while moving and playing the game
Ignores occlusion	User ignores the fact that virtual objects are not occluded by physical objects
Ignores misalignment	User ignores misalignment between virtual items and previous physical positions
Problem Severity	
1 (low)	Child could continue without help.
2 (medium)	After being helped, child continued.
3 (high)	Facilitator took over, or child wanted to stop.

3.4 AUGMENTED REALITY FOR EDUCATION

I have investigated augmented reality's applications to education through several activities, beyond the educational systems described above. In several theory-driven investigations, my colleagues and I have argued that AR has specific affordances that make it beneficial to learning. These affordances include the ability to reduce a learner's cognitive load by directing their attention to important pieces of the physical environment, present multiple representations of the learning content, providing avenues for student collaboration while within 3D simulations, and appealing to multimodal learning and engaging embodied cognition (Radu, Zheng et al. 2010, Bujak, Radu et al. 2013, Radu 2014). Through a meta review of existing studies of AR-based learning, I have identified the educational domains where AR has been shown to have benefits, and I have discussed how the learning benefits differ between AR and other media such as textbooks, computer-based learning, and virtual reality (Radu 2012, Radu 2014). Furthermore, during my employment at PBS KIDS, I have directed the design, research and production of the tablet-based AR game Cyberchase Shape Quest, and this work has generated a handheld AR geometry game, a process for designing handheld AR games for young children, and a set of usability issues encountered by 6-8 year old children while using handheld augmented reality (Radu, Doherty et al. 2015). Finally, through collaboration with PBS KIDS and WestEd, we have investigated the difficulties which Grade 1-3 teachers have in teaching the Common Core State Standards mathematics curriculum, and designed AR prototypes to address problematic topics such as fractions, number decomposition and word problems (Radu, McCarthy et al. 2016).

Through these experiences, I have discovered that usability is a crucial component of designing educational products for young children, because, in order to maximize learning, children must remain motivated and engaged with the product. Handheld augmented reality games are quite different than traditional handheld games, and no guidelines currently exist for designing AR applications for young children. Thus design experts typically have difficulties predicting how children will react to specific AR designs: what kinds of designs will be too difficult for an age group, which designs will be challenging enough to be motivating, and how old children will need to be in order to use a specific AR design. Because of these issues, the process of designing AR games is prohibitive as there is significant amount of uncertainty during game design phases (both regarding children's ability to use the design, as well as understanding what AR designs will lead to improved learning), and because early and frequent user testing is required during implementation phases. Combined with the comparatively high cost of producing handheld-AR applications, this lack of knowledge is currently a major issue restraining the commercial production of children's educational AR applications. Knowledge of usability of handheld-AR applications would therefore become a valuable asset for educational technology designers.

CHAPTER 4

EXPERIMENTAL STUDY OF RELATIONSHIPS BETWEEN YOUNG CHILDREN'S AGE, PERFORMANCE, AND USABILITY ISSUES

4.1 OVERVIEW OF THE ARC STUDY

The following study was motivated by an interest in understanding differences in AR usability across age groups, and interest in understanding whether specific types of AR interaction techniques are child-appropriate. Based on the previously described work, my hypotheses are that when young children are exposed to AR experiences, they exhibit slower performance and more inaccurate interactions, as well as more types of usability issues, than compared to older children. Furthermore, children are expected to exhibit different performance and usability issues when they are exposed to different kinds of interaction techniques, and the performance and usability issues are expected to vary according to the complexity of the interaction, on the dimensions of "Selection Type" (Finger vs. Crosshair, whereby the former interaction requires independent hand movements and the latter does not) and "Movement Type" (Tunnels vs. No Tunnels, whereby the former requires children to perform full-body movements around the gameboard). These topics have not been investigated by previous work, and will be the focus of my study. In this study, I am interested in answering the following main research questions. Table 1.1.1 presents more details of sub-questions and hypotheses, while Table A.1 presents additional summary of the methods employed to answer these questions.

RQ1: How does children's age impact performance and usability issues in handheld-AR interactions?

RQ2: How do different handheld-AR interaction techniques compare, in terms of performance and usability issues encountered by children?

RQ3: What types of usability issues are experienced by children in handheld-AR?

To investigate the above research questions, I designed a mixed-methods experiment, which used both quantitative and qualitative methods to measure children and their abilities to use augmented reality. The main research instrument was an augmented reality game designed for children 5-10 years old, and containing several levels, with each level testing a different interaction technique condition. In the following sections, I describe the design of this study. First, I will discuss the reasons for selecting specific interaction techniques and age groups. Then, I will describe the game and its validation through pilot studies. This will be followed by quantitative and qualitative metrics, followed by the data analysis and results.

4.2 GAME AND EXPERIMENTAL DESIGN

4.2.1 Game Structure and Narrative

The game is structured as a typical tabletop augmented reality game, where a three-dimensional virtual world appears on top of the "gameboard" paper once it is viewed through a smartphone camera (Figure 4.2.1). In the experimental setup, the game was placed on a table without chairs, and children had space to move around the table at their leisure. The height of the gameboard was adjusted to be at the level of each player's stomach, such that each player could comfortably observe the gameboard through the smartphone.

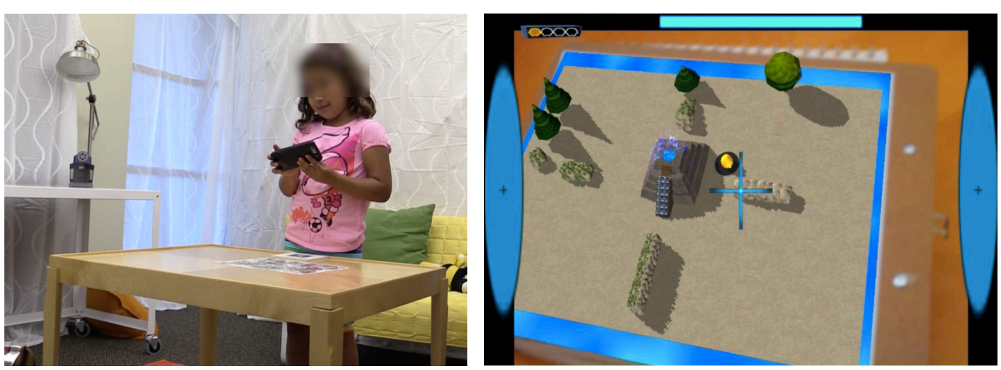


Figure 4.2.1 Child playing the game, in a level where lemons were collected through Crosshair Selection.

In the game, the player took the role of a wizard who must collect a set of magical lemons in order to create items for their pets. The game was composed of a tutorial plus 4 levels. Each of the 4 main levels of the game was associated with a different AR interaction technique condition, and was composed of a “lemon collection” phase and a “mini game” phase. During lemon collection, the player had to gather 4 sets of 4 lemons each, for a total of 16 lemons. Once all lemons were collected, they magically transformed into a play item and the mini game started. In the mini game, the child played with their pet and the newly-created item (Figure 4.2.2).

During the “lemon collection” phase, all children played the augmented reality game while standing or moving around the table. During the “mini game” phase at the end of each level, the game did not involve any augmented reality and did not require children to look at the gameboard; thus children were asked to sit during this phase. The mini game was a requirement of the game design, in order to give children a rest period from standing while at the same time offering entertainment and agency. During the data analysis, only data from the lemon collection phases was analyzed.

Figure 4.2.2. The game characters and magical objects (left), and the mini-game associated with one completed object (right).

Prior to the set of 4 gameplay levels, children were exposed to a tutorial phase. During pilot testing I determined that children were performing poorly because they did not have previous exposure to the technology and generally did not understand that the game should be played by moving the body in the physical space. Thus, a tutorial was developed to familiarize children with augmented reality technology. The tutorial was designed to give the player an introduction to augmented reality (requiring children to explore the 3D game world by viewing it from different angles and moving their body around the physical space), and also giving the player an introduction to the game mechanics and all interaction techniques used in subsequent levels of the game.

4.2.2 Interaction Technique Factors

In this study, I investigate two different selection techniques, under two movement conditions (Figure 4.2.3). These variations were driven by two factors: Selection Type and Movement Difficulty.

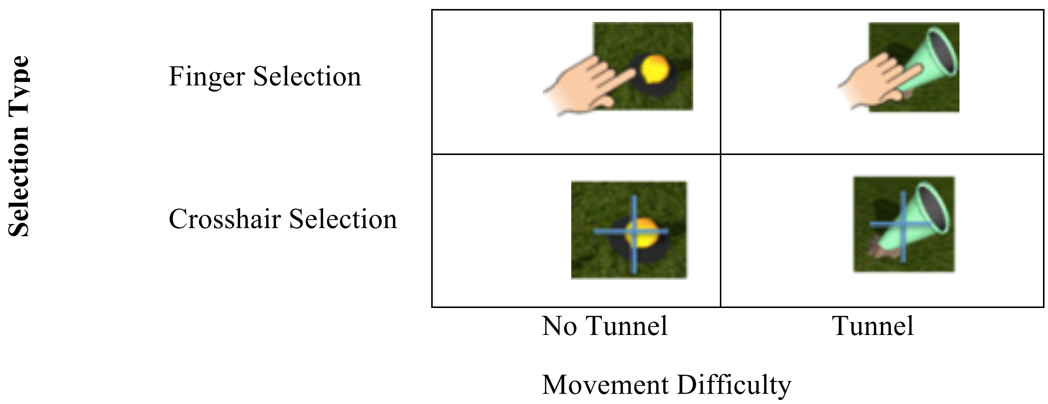


Figure 4.2.3. The interaction conditions tested in the ARC study.

There were two Selection Types: Finger Selection or Crosshair Selection. This dimension indicates whether children need to perform independent movements with each of their hands. These selection types have been chosen because they represent the most popular interaction techniques in current handheld AR games for young children. When using Finger Selection, players must touch their finger to the screen position where a target lemon appears. This interaction technique requires independent hand movement, as users must touch the screen at specific locations while holding the device steady. In Crosshair Selection, the players have a crosshair in the center of their screen, and they select a target lemon by touching one of the screen side buttons once the lemon is overlapping the crosshair (Figure 4.2.1 right). This technique can be performed while both hands are holding the device, and does not require users to touch specific screen locations; it does, however, require that the user precisely position the center of the screen. Each selection type has an accuracy distance threshold, which was empirically determined. In the Finger Selection mode, the lemon surface can be 60 pixels away from the screen touch point; this corresponded to 4.55mm in our experimental smartphone. This threshold was selected to account for the problem of finger occlusion (Boring, Ledo et al. 2012), and it corresponds to the size of an average 3-10-year old child's finger (Hohendorff, Weidermann et al. 2010). In Crosshair Selection, the lemon can be 35 pixels away from the center of the crosshair (2.7 mm). This distance was used to account for angular reorientation of the phone while children press the crosshair buttons, and was determined during our pilot study by analyzing how much children shake the phone in the last 1s prior to touching the crosshair buttons.

Movement Difficulty, the other factor influencing our interaction conditions, impacts whether the players need to change perspective during gameplay. It has two types: No Tunnels and Tunnels. In the No Tunnels condition, the player can see the targets from any angle and does not need to move their body. In the Tunnel condition, the targets are enclosed in virtual tunnels, thus the player must change their perspective in order to select the target. The targets in the Tunnel conditions were rotated such that, between each lemon, players were forced to change their angle relative to the gameboard by 45 degrees while remaining relatively in front of the table. This variation has been chosen because it covers the current and future state of AR games: currently, most handheld AR games do not require the user to understand 3D spatial perspectives or change their point of view (see Chapter 2.3.1), but as AR applications become more spatially complex, it is expected that AR games will require users to understand spatial relationships and change their point of view.

Each child was exposed to 4 different interaction technique conditions, varying on the conditions of Selection Type and Movement Difficulty (described above). The Selection Type conditions were randomized between players, while the Movement Difficulty conditions were not randomized (all players experienced No Tunnels before Tunnels). The game environment was randomized between conditions.

4.2.3 Age Groups

The age of the participants was chosen to be 5 to 10 years old. This age range has been chosen for several reasons. First, it spans across significant changes identified by developmental psychology, spanning across the 7-years old boundary of Piaget's "preoperational" and "concrete operational" stages, and covering significant developments in physical and cognitive development (Chapter 2). It is worth noting that researchers have criticized the specific age boundaries proposed by Piaget, and it is not expected that children will show a radical change in skill at the boundary age of 7. Research criticizing Piaget has shown high variability in children, specifically that children within the same age group do not all possess the same cognitive capabilities; and, depending on experimental conditions, children can show some cognitive competencies at an earlier age than Piaget's stages predict (Flavell, Beilin et al. 1992, Rosser 1994, Fischer and Immordino-Yang 2002, Thornton 2002, Kesselring and Müller 2011). Nevertheless, Piaget's stages can be useful for illustrating the general changes that occur in children's cognition and for providing a set of initial predictions about the variety of usability issues that children may encounter with augmented reality (see Chapter 2 for more details).

Second, this age range represents a fruitful area for AR educational applications appealing to children in Kindergarten to Grade 4, and spanning ages in which children shift from playing with concrete toys to understanding abstract concepts (Bujak, Radu et al. 2013, Radu, McCarthy et al. 2016). Additionally, research in augmented reality has not been focused on designing handheld augmented reality for children as young as 5 year olds, and my discussions with child technology designers have revealed a reluctance to design AR for such young children, due to potential limitations in children's skills and the limited design knowledge on the topic. However, in my previous studies I have observed that children as young as 5 years old were capable of interacting with handheld augmented reality when properly trained, thus it possible for such young children to use this technology. Therefore, the age range of 5 to 10 years old was selected for this investigation.

4.2.4 Game Software Architecture

The game was implemented in the Unity3D engine (Unity Technologies 2016) and deployed to a Motorola Atrix 2 smartphone during the user studies. The game uses the Vuforia augmented reality library (PTC Inc 2016) in order to track an image printed on flat letter-size paper. The game software architecture (Figure 4.2.4) is designed to allow full replay of a user's session after the participant completes the experiment. While participants play the game during the experiment, the game logs all the player movements and interactions. After the study completed, the participants' game sessions can be fully simulated and analyzed based on the log files. This architecture allows experimenters to have complete access to a user's game session, in order to view the user's screen without the need for processor-intensive live screen recording software, and in order to answer questions that may arise after the experiment has completed.

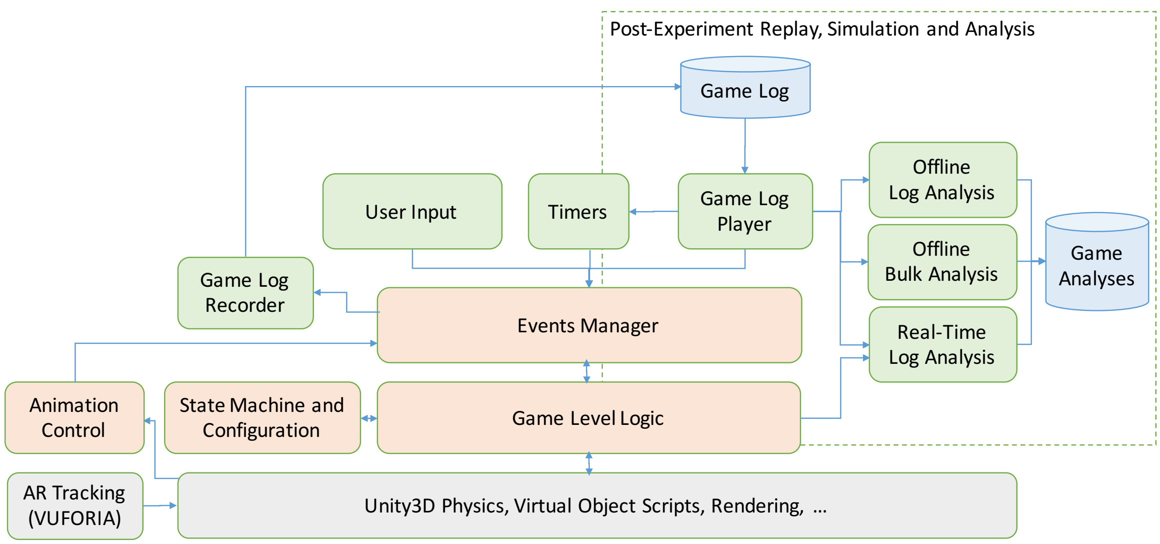


Figure 4.2.4. Software architecture of the ARC game.

4.3 STUDY DESIGN, VARIABLES AND METRICS

The research questions and hypotheses associated with each question are outlined in **Error! Reference source not found.**. In order to investigate these research questions, the study has been designed as a mixed-methods study with between-subjects and within-subjects variables.

The within-subjects factors that were manipulated were Selection Type and Movement Difficulty. The between-subjects factor was children's Age. Analysis was performed using both quantitative and qualitative methods. The following section describes the independent and dependent variables in this study.

Table 4.3.1 Variables investigated in the ARC study.

- **Independent variables related to Interaction Technique**
  - Selection Type (Nominal 2 types: Finger Selection, Crosshair Selection)
  - Movement Difficulty (Nominal 2 types: No Tunnels, Tunnels)

- **Independent variables related to Physical and Cognitive Development**
  - Age (Ordinal 3 groups: 5-6, 7-8, 9-10 years old)
  - Hand Size (Continuous)
  - Woodcock-Johnson-III Spatial Relations Test (Continuous) (Woodcock 1990)
  - NEPSY-II Visuomotor Precision Test (Continuous) (Kemp, Korkman et al. 2001)
  - NEPSY-II Block Construction Test (Kemp, Korkman et al. 2001)

- **Dependent variables related to Performance (from software logs)**
  - Task Completion Time (s) (Continuous)
  - Number of Tracking Losses (Continuous)
  - Time to Recover Tracking (s) (Continuous)
  - Number of Selection Errors (Continuous)

- **Dependent variables related to Subjective Ratings (from post-survey)**
  - Ease of use (Ordinal)
  - Fun (Ordinal)
  - Comfort (Ordinal)

- **Dependent variables related to Observed Usability Problems (from video observations)**
  - Types of Observed Usability problems (Nominal)
  - Severity of Observed Usability Problems (Ordinal)

- **Other independent variables (collected in questionnaires)**
  - Gender (2 groups: Male, Female) (Nominal)
  - Previous experience with technology (Continuous)

4.3.1 Research Variables

This section will discuss the variables involved in the study, listed in Table 4.3.1.

4.3.1.1 Experimentally Manipulated Variables

Two variables will be experimentally manipulated in the augmented reality game: Selection Type and Movement Difficulty. The variable of Selection Type determines if the child will be selecting virtual objects by tapping their finger on the screen or by using a crosshair with side buttons. The Movement Difficulty determines if the child can play the game by staying still or if they need to significantly move around the gameboard. More information about the conditions can be found in Section 4.2. The combination of variables yields a 2x2 design, resulting in 4 conditions each corresponding to a level of the augmented reality game.

4.3.1.2 Physical and Cognitive Development Factors

In this study the children's *Gender* were both Male and Female, and their *Age* was between 5 and 10 years old. For analysis I measured participant ages in months, and also grouped age into three groups: 5-6, 7-8, and 9-10 year olds.

I measured participants' *Hand Size* because it should influence how easily children are able to manipulate the phone and perform various interaction techniques.

Children were exposed to three kinds of developmental tests, from the NEPSY II and Woodcock Johnson III batteries:

Spatial Relations Test – this test asks children to solve two-dimensional spatial puzzles. To receive a high score, children must be able to isolate shapes and perform mental rotation tasks.

Visuomotor Precision Test – in this activity, children are asked to follow a path while using a pencil. To receive a high score, children must employ visuomotor (hand-eye coordination) skills.

Block Construction Test – in this activity, children are shown a figure of a three-dimensional block structure and asked to build the structure using physical toy blocks. To receive a high score, children must employ visuospatial reasoning and perform fine-motor physical manipulation.

4.3.1.3 *Performance Measures*

To measure children's reaction to the augmented reality experience I used performance measures collected through the software. These measures were calculated for each of the 4 game conditions, by averaging the 16 lemons collected by a participant within each condition. They are as follows:

Task Completion Time: The average amount of gameplay time (in seconds) a child spent in order to collect each lemon. This metric does not include time spent during tracking loss.

Number of Tracking Losses: The average number of times that the child lost tracking. Tracking loss occurs when the phone camera stops seeing the gameboard image.

Time to Recover Tracking: The average amount of time that a player needed in order to exit tracking loss and resume playing. This amount of time was not included in the Task Completion Time measure, and was calculated only when Number of Tracking Losses was non-zero.

Number of Selection Errors: The average number of invalid selections performed while attempting to collect a lemon. In the Finger Selection condition, an invalid selection occurs when the player touches the screen outside the target area. In the Crosshair Selection this occurs when the player clicks the selection button, but the crosshair is pointing away from the target area.

4.3.1.4 *Subjective Measures*

At the end of each level I asked children to report how they felt, by asking if they were Comfortable, if the level was Easy, and if they had Fun while playing. The questions were randomized between levels, and delivered using a modified rating scale (Figure 4.3.1) based on the Smileyometer scale [39].

Figure 4.3.1. Example subjective experience questionnaire item.

4.3.1.5 Usability Problems

Children were observed playing the augmented reality game, and usability problems were detected through qualitative analysis. The analysis identified a total of 16 kinds of problems, and each problem was associated with a severity. Section 4.4.6 will describe each problem, and the process of identifying usability problems through video coding and qualitative analysis.

4.3.1.6 Other variables

Other variables of interest collected through the questionnaire or observations were: Child's Gender (Male or Female) and Previous Experience with Technology.

4.4 METHODS AND RESULTS

4.4.1 Participants and Study Protocol

Children were recruited from both Georgia Institute of Technology and Emory University. Convenience-sampling methods were used in both cases. On the Georgia Tech campus, parents responded to paper-based recruitment fliers placed across campus. The Emory University's psychology department child subject pool was also used to contact families via phone or email. In all cases, families were told about the opportunity to participate in a study that investigates smartphone-based games for children. Upon arrival at our study lab, parents and children were informed about all the study procedures and asked if they are comfortable participating; voluntary participation was recorded by signing our study's IRB-approved consent forms.

The study was composed of 4 phases: pre-game interview phase; gameplay phase; post-game interview phase; and developmental tests phase. The materials used for data collection in these phases is included in Appendix D.

At the beginning of the study, questions were administered to parents through a study questionnaire, and to children through a structured interview with the experimenter. The parent questionnaire asked about the child's date of birth, experiences with technology, arts and crafts, play habits, and any restrictions from technology use. The child pre-study interview was designed to determine a child's experiences with technology similar to augmented reality, and was also intended to build a rapport between the experimenter and child participant prior to the main portions of the study.

After the study pre-interview, children played with the smartphone game described in Section 4.2.1. The game was composed of a tutorial phase, followed by 4 levels of gameplay, each corresponding to an experimental condition. Within each level, there was a "lemon collection" phase, followed by a mini-interview, followed by a "minigame" phase. Each game level corresponded to one of the four experimental conditions. The conditions varied on the variables of Selection Type and Movement Difficulty described above. The Selection Type conditions were randomized between players (some players experienced the Crosshair condition before Finger; while others experienced the Finger condition before Crosshair), while the Movement Difficulty conditions were not randomized (all players experienced No Tunnels before Tunnels). Within each game level, the player collected 4 sets of 4 lemons, for a total of 16 lemons per level. After the lemon collection, the game loaded a min-game phase, where the child could relax and play a short fun game not related to augmented reality or lemon collection. The mini-game was designed to take a few minutes before it was fully loaded. This was specifically designed in order to allow time for a mini interview between child and experimenter after each level. The mini-interview asked children questions about the degree to which the recently-played game level was enjoyable, difficult, and comfortable (see Appendix D). After all questions were answered, the child played with the mini-game before proceeding to the next game level.

After all the levels in the game were completed, a post-game interview was administered. The structured interview asked questions such as: if the child had played similar games before, how he/she felt about having to move around a gameboard while playing a game, which parts of the game were favorite and least favorite, whether he/she would share the game with friends, and asked children to again rate their enjoyment / comfort / difficulty over all the game levels. This concluded the research study phase related to playing with the augmented reality game.

For the final phase of the experiment, the child, parent and experimenter moved into a new room environment, where children performed standardized developmental tests. The test batteries were described in Section 4.3.1.2 and included a spatial relations test, fine motor test, and block construction test.

4.4.2 Data Collection

Data was collected in a variety of mediums including gameplay logs, video recordings and paper-based recordings. The link between experimental variables and data collection medium is listed in Section 4.3.1.

Based on the power analysis from the pilot study, I determined that 16 participants per age group (total 48 participants) would be satisfactory for finding significant main effects for each of the independent variables and for identifying several significant correlations to developmental skills. After the study data collection completed, 3 children were omitted (1 child decided to quit the game in the middle of gameplay, 1 child did not finish the study because the parent had to leave, and 1 child experienced technical issues with the game logging).

After data collection was completed, I performed outlier removal based on metrics of children's performance and developmental tests. During some trials of individual lemon collection, I observed participants stopping their gameplay due to extraneous events, such as stopping to clear runny noses, interrupting the gameplay to say something to the experimenter, etc. These interruptions would impact the length of time a child took in order to collect an individual lemon and impact the average amount of time for all trials; thus for the purposes of data analysis I excluded individual trials in which a lemon completion time was beyond 2.5 standard deviations of a child's average times within the same experimental condition. This accounted for 3% of all trials.

During the experiment, I observed that a few children took a long time overall to complete the lemon-collection task compared to other children. Such outliers can indicate the presence of general problems that significantly influence children's ability to interact with AR games. To find such outliers, I identified children whose Task Completion Time scores were beyond 2.5 standard deviations of the mean within their age group. This accounted for 3 children removed from the final analysis.

Furthermore, I attempted to identify children whose age-standardized developmental test scores were beyond 2.5 standard deviations past the mean within each age group. Such outliers might indicate the presence of child development issues. No children matched this criterion.

The resulting dataset consists of 42 children of both genders across three age groups (Table 4.4.1.). All the analysis presented below is based on this dataset, with the exception of the Observed Usability Problems analysis. For the Observed Usability Problems analysis, 2 children were excluded since video recordings were not captured due to technical issues on the experiment day (these children were both age 10 years old).

Table 4.4.1. Demographics of the children in the ARC study.

Age Group	Females	Males	Total
5-6 year olds	6	8	14
7-8 year olds	7	7	14
9-10 year olds	7	7	14

4.4.3 Analysis of Performance Metrics

Please refer to Appendix C for descriptive statistics related to the variables involved in this analysis. The performance metric analysis focused on the effects of the variables of Age Group, Selection Type and Movement Difficulty as they influence the dependent measures of Task Completion Time, Number of Tracking Losses, Time to Recover Tracking, and Number of Selection Errors. Generally, for each dependent measure the analysis was performed using a repeated-measures ANOVA, with Selection Type and Movement Difficulty as the within-subjects factors, and Age Group as the between-subjects factor. Significant between-factor effects were followed-up with post-hoc Tukey tests; significant interaction effects were followed using contrasts and Bonferroni alpha corrections.

If assumptions of the ANOVA test were violated, I attempted to transform the data such as not to violate the assumptions. If a suitable data transformation was not found, I performed nonparametric Kruskal-Wallis H tests (for the between-subjects factor) or Wilcoxon signed-rank tests (for the within-subjects factors), with Bonferroni-corrected Type I thresholds. Whenever the ANOVA assumptions were violated but nonparametric tests showed the same significant differences between groups, I report the ANOVA results.

4.4.3.1 Task Completion Time

When analyzing the dependent measure Task Completion Time, the data violated the normality assumption of the parametric test. To meet the assumptions, the data was transformed using a reciprocal transformation. The parametric analysis on the transformed data yielded the same significant results as the original analysis, thus the original analysis is reported below.

A significant interaction effect between Movement Difficulty and Age was found ($F(2,39)=3.59$, $p=0.037$). This effect was analyzed first because it can point at differences within specific age groups or within specific movement difficulties. I performed 5 post-hoc Bonferroni-corrected contrasts to investigate the interaction effect, while keeping the Type I error threshold at 0.01. First, I investigated whether there are statistically significant differences between Age Groups within each of the two Movement Difficulty conditions; such **statistically significant differences were found between Age Groups for the No Tunnel** ($F(2,39)=25.44$, $p<0.001$) **and Tunnel** ($F(2,39)=16.23$, $p<0.001$) conditions. Post-hoc analysis indicated that in both cases, the mean task completion time of 5-6 years-old children was significantly slower than compared to the 7-8 years-old children and compared to the 9-10 years-old children ($p<0.001$ in all cases). In the case of no tunnels, the average time 5-6 year olds took to collect one lemon ($M=3.0s$, $SD=0.7s$) was 40% slower than for 7-8 year olds, and 37% slower than for 9-10 year olds. Similarly, in the case of tunnels, the average time 5-6 year olds took to collect one lemon ($M=6.1s$, $SD=1.6s$) was 31% slower than for 7-8 year olds, and 33% slower than for 9-10 year olds. However, the mean task completion time of 7-8 years-old children was not significantly different than that of the 9-10 years-old children. Finally, I investigated whether the differences between No Tunnel vs. Tunnel conditions differed within each of the three Age Groups. **Within each Age Group, the differences between No Tunnel and Tunnel conditions was statistically**

significant (p<0.001), with Tunnel conditions showing slower completion times across all Age Group. For 5-6 year olds, collecting the average lemon in tunnel levels (M=6.1s, SD=1.6s) was 51% slower than for non-tunnel levels; for 7-8 year olds, collecting a lemon in tunnel levels (M=4.2s, SD=0.61s) was 56% slower than non-tunnel levels; finally, for 9-10 year olds, collecting a lemon in tunnel levels (M=4.1s, SD=0.7s) was 53% slower than non-tunnel levels. This study of the interaction effects was consistent with the significant main effects reported below; thus it is unclear between what specific conditions there is an interaction between Age Group and Movement Difficulty.

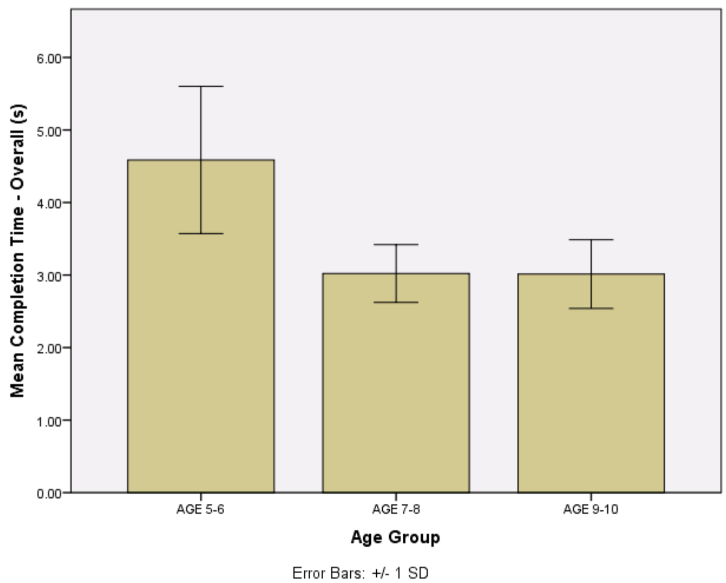


Figure 4.4.1. Average Task Completion Time (measured in seconds per Lemon) vs Age Group.

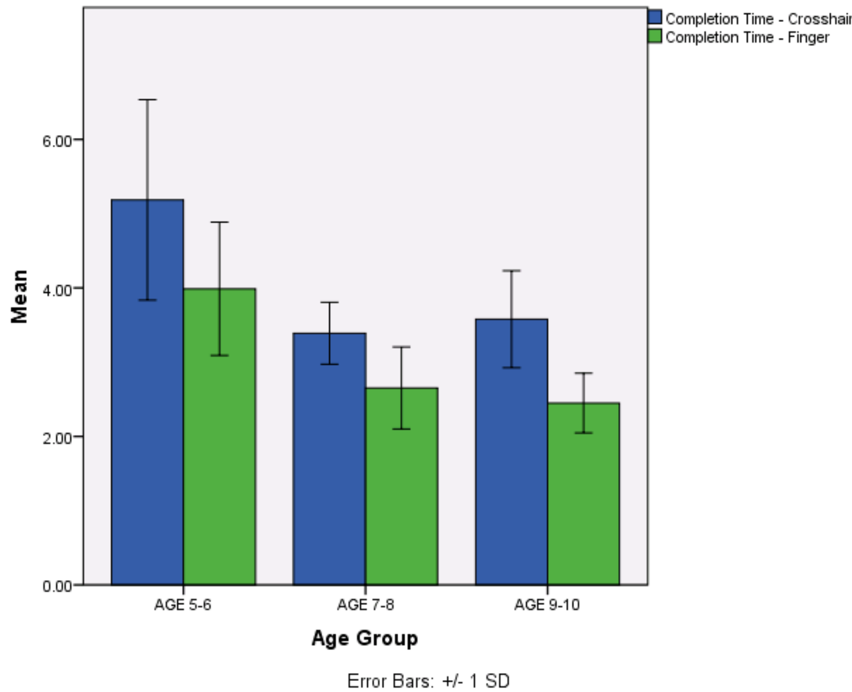


Figure 4.4.2. Average Task Completion Time (measured in seconds per Lemon) vs Selection Types.

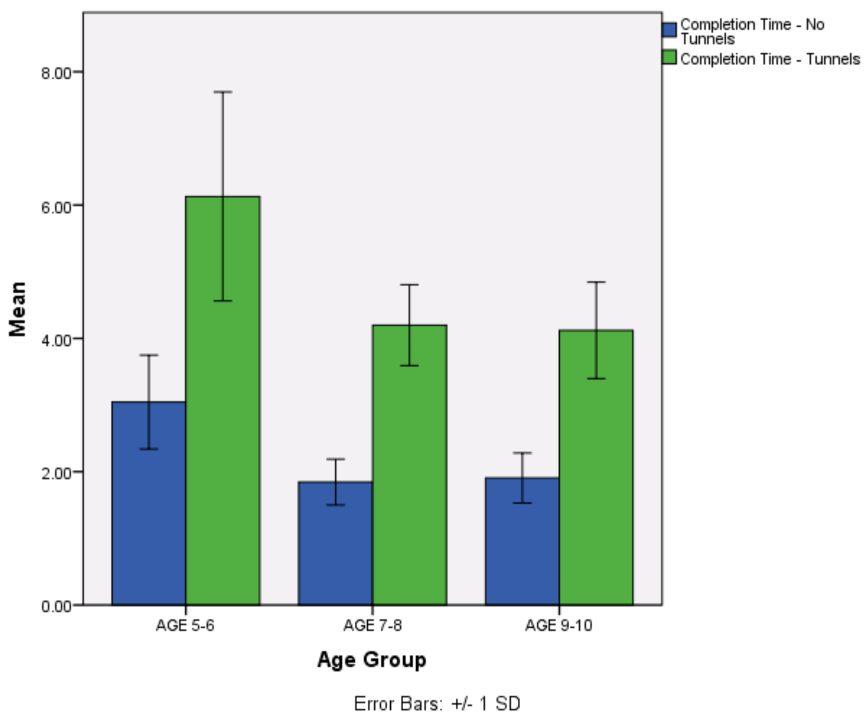


Figure 4.4.3. Average Task Completion Time (measured in seconds per Lemon) vs Movement Difficulties.

A significant main effect of Age Group was found (F(2,39)=24.39, p<0.001). Post-hoc analysis indicated that, between all gameplay conditions, the mean task completion time of 5-6 years-old children (M=4.59s, SD=0.183s) was significantly slower than compared to the 7-8 years-old children, and compared to the 9-10 years-old children by an average 34% in both cases; however, the mean task completion time of 7-8 years-old children was not significantly different than that of the 9-10 years-old children. **A significant main effect of Selection Type was also found** (F(1,39)=76.05, p<0.001). Overall, selecting an item with the Finger (M=3.03s, SD=0.10s) was significantly faster than the Crosshair, by an average 25%. **A significant main effect of Movement Difficulty was also found** (F(1,39)=321.68, p<0.001); selecting an item in the No Tunnel levels (M=2.27s, SD=0.07s) was overall faster than Tunnel levels by an average 53%.

Overall, children's age had a significant overall effect on completion time, and children in the 5-6 years old group were significantly slower than the 7-8 and 9-10 year olds; no significant differences were found between the 7-8 and 9-10 year olds. Selecting items in the Finger condition was faster than in the Crosshair conditions. Selecting items in the No Tunnel conditions was significantly faster than in the Tunnel conditions.

For the descriptive statistics related to this metric, please see Appendix C. The data informs the research questions as follows:

RQ1-1: Does speed of selection differ between age groups?

***Hypothesis**: Younger children will be slower at performing selection tasks.*

***Results**: Across all movement conditions, 5-6 years old group were significantly slower than the 7-8 and 9-10 year olds; however, no significant differences were found between the 7-8 and 9-10 year olds*

RQ2-1: Does speed of selection differ between interaction techniques?

***Hypothesis**: Interaction techniques that involve independent hand movements will lead to lower speed.*

***Results**: The opposite result was found - selecting items in the Finger condition (independent hand movement) was significantly faster than in the Crosshair conditions.*

__Hypothesis2__: Interaction techniques that involve whole-body movement will lead to lower speed.

__Results__: Selecting items in the Tunnel conditions (whole body movement) was significantly faster than the No Tunnel conditions.

4.4.3.2 Number of Tracking Losses

When analyzing the dependent measure Number of Tracking Losses, the data violated the normality assumption of the parametric test. Nonparametric analysis yielded the same significant results as the parametric analysis. The parametric analysis is reported below.

A significant interaction effect was found between Movement Difficulty and Age ($F(2,39)=5.48$, $p<0.01$). This effect was analyzed first because it can point at differences within specific age groups or within specific movement difficulties. I performed 5 post-hoc Bonferroni-corrected contrasts to investigate the interaction effect, while keeping the Type I error threshold at 0.01. First, I investigated whether there are statistically significant differences between Age Groups within each of the two Movement Difficulty conditions; such **statistically significant differences were found between Age Groups for the Tunnel conditions** ($F(2,39)=6.63$, $p=0.003$), but **no statistically significant differences were found between Age Groups for the No Tunnel conditions** ($F(2,39)=0.160$). This indicates the number of tracking losses may not be different between age groups for the No Tunnel conditions ($M=0.03$, $SD=0.05$ per lemon collected). For the Tunnel conditions, post-hoc analysis indicated that the mean number of tracking losses encountered by 5-6 years-old children ($M=0.21$, $SD=0.12$) was on average 59% higher than for 7-8 years-old children ($p=0.017$) and 68% higher compared to the 9-10 years-old children ($p=0.005$); however, the mean number of tracking losses of 7-8 years-old children was not significantly different than that of the 9-10 years-old children. Finally, I investigated whether the differences between No Tunnel vs. Tunnel conditions differed within each of the three Age Groups, in terms of number of tracking losses. Within the 5-6 year old group, **the differences between No Tunnel and Tunnel conditions was statistically significant** ($p<0.001$), with tunnel levels showing 77% more tracking losses than no-tunnel levels, and **similarly significant differences were detected within the 9-10 year old group** ($p=0.008$), with Tunnel conditions showing 66% less tracking losses than No Tunnel levels. However, **within the middle age group, the 7-8 year olds, the number of tracking losses was marginally not statistically significant** (p=0.031 – marginally not significant) beyond the Type I error threshold set for contrasts, possibly indicating that for 7-8 year olds the number of tracking losses are not different between No Tunnels vs. Tunnels conditions.

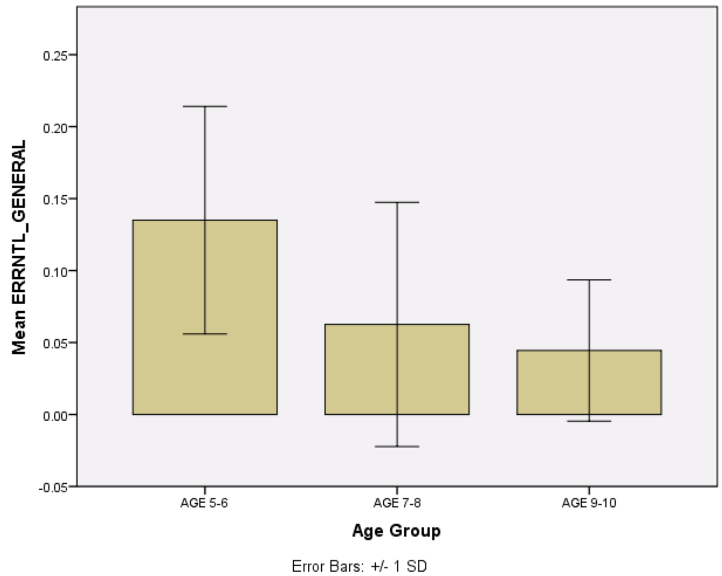


Figure 4.4.4. Average Tracking Losses (per Lemon) vs Age Group.

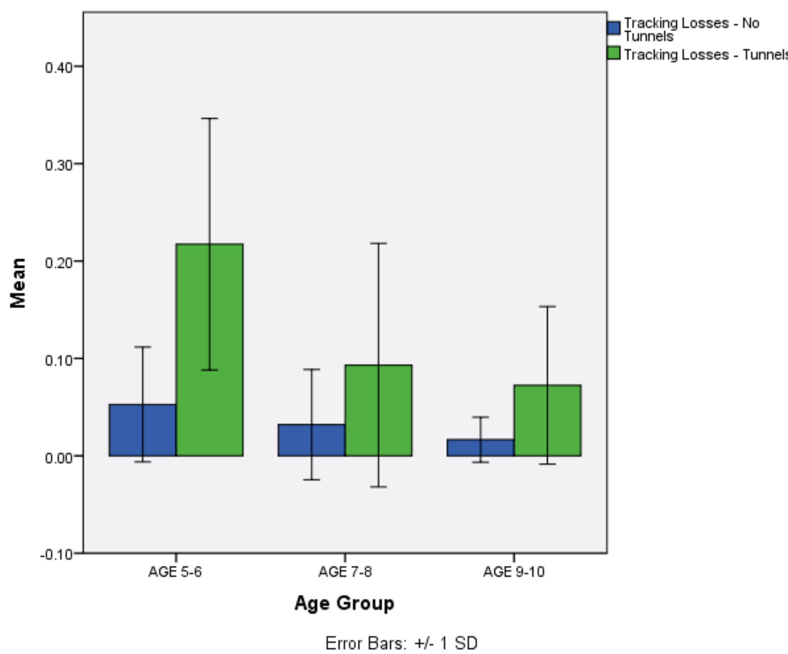


Figure 4.4.5. Average Tracking Losses (per Lemon) vs Movement Difficulty.

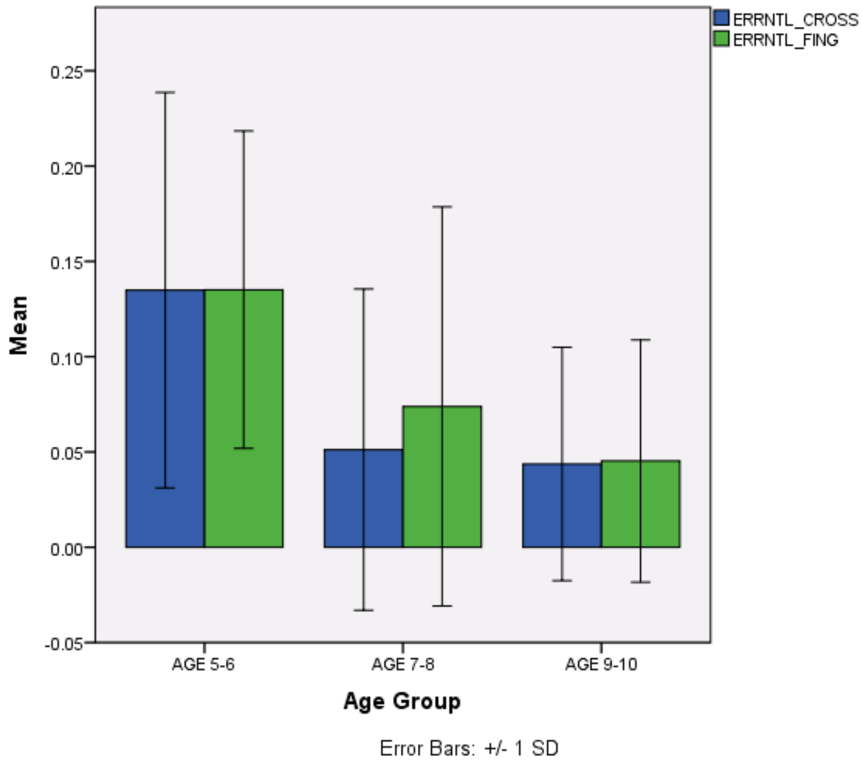


Figure 4.4.6. Average Tracking Losses (per Lemon) vs Selection Type (Crosshair or Finger).

A significant main effect of Age Group was found overall (F(2,39)=6.09, p<0.01). Post-hoc analysis indicated that the mean number of tracking losses of 5-6 years-old children (M=0.1, SD=0.08) was significantly higher by an average 54% compared to the 7-8 years-old children, and higher by an average 67% compared to the 9-10 years-old children; however, the mean number of tracking losses of 7-8 years-old children was not significantly different than that of the 9-10 years-old children when considering all conditions. **A significant main effect of Movement Difficulty was also found** (F(1,39)=38.52, p<0.001), indicating that overall the Tunnel conditions generate more tracking losses (M=0.1, SD=0.1 per lemon) by an average 73% than compared to No Tunnel conditions. **No significant main effect of Selection Type was found** (F(1,39)=0.354, p=0.56); overall, losing tracking was not significantly more frequent between Finger vs. Crosshair conditions.

This analysis indicates that, overall, 5-6 year olds perform significantly worse than 7-8 and 9-10 year olds; while 7-8 and 9-10 year old groups are statistically similar to each other. Looking within the Movement Difficulty factor, it can be observed that this age effect holds for Tunnel conditions but not for No Tunnel conditions; in the Tunnel conditions the 5-6 year olds are significantly different, but in the No Tunnel conditions no effect of age was detected. Furthermore, within each age group the analysis indicates that Tunnel conditions cause significantly more tracking losses than No Tunnel conditions within 5-6 and 9-10 year old groups; but this effect is not significant for 7-8 year olds. The factor of Selection Type was not statistically significant, thus indicating that the number of tracking losses within Finger vs. Crosshair conditions may not be different.

For the descriptive statistics related to this metric, please see Appendix C. The data informs the research questions as follows:

RQ1-3: Does accuracy for AR tracking differ between age groups?

> ***Hypothesis****: Younger children will have a higher frequency of AR tracking losses.*
>
> ***Results****: When selecting items in conditions requiring whole-body movement (tunnel levels), 5-6 year olds encountered significantly more tracking losses than 7-8 and 9-10 year olds; no significant differences were found between 7-8 and 9-10 year olds. When selecting items in conditions not involving whole-body movement (no-tunnel levels), there were no significant differences found between age groups.*

RQ2-3: Does accuracy for AR tracking differ between interaction techniques?

> ***Hypothesis1****: Interaction techniques that involve independent hand movements will lead to higher frequency of tracking losses.*
>
> ***Results****: No significant difference was found between interactions involving independent movements (Finger selection) vs. coordinated movements (Crosshair selection).*

***Hypothesis2**: Interaction techniques that involve whole-body movement will lead to higher frequency of tracking losses.*

***Results**: Overall, interaction techniques that involve whole-body movement (Tunnel levels) led to significantly higher frequency of tracking losses than non-whole-body movement conditions (No Tunnel levels). For 5-6 and 9-10 year old children, such a difference between conditions was statistically significant; for 7-8 year olds, the difference between conditions was marginally not significant.*

4.4.3.3 Time to Recover Tracking

When analyzing the dependent measure Time to Recover Tracking, the data violated the normality assumption of the parametric test. Nonparametric analysis yielded the same significant results as the parametric analysis. The parametric analysis is reported below.

The dependent measure of Time to Recover Tracking was generated only for conditions where a child had experienced tracking losses. Only 4 children had tracking losses in all of the four gameplay conditions at once, and none of the children were in the 9-10 years old group; therefore, I decided that a repeated-measures ANOVA analysis using the 4 within-subjects conditions would be ineffective for investigating the effects of Age Group, Movement Difficulty, or Selection Type variables. Instead I performed 3 separate tests with Type I error set at 0.016. First, a one-way ANOVA was performed to determine the effects of between-subjects factor Age Group, on children's time to recover tracking averaged over all Movement Difficulty and Selection Type conditions. **A significant main effect of Age Group was found** ($F(2,30)=7.67$, $p=0.002$); post-hoc Tukey tests show that 5-6 year old children ($M=2.3s$, $SD=0.9s$) are on average 52% slower at recovering tracking than compared to both 7-8 year olds, and on average 55% slower compared 9-10 year olds ($p<0.01$), but 7-8 and 9-10 year olds were not statistically different in terms of overall time to recover tracking. The final two tests were paired T-tests on children's average time to recover tracking, indicating **no significant mean differences between Finger vs. Crosshair conditions** ($p=0.643$); and **no significant mean differences between Tunnel vs. No Tunnel conditions** ($p=0.279$). **No interaction effects** between variables was found.

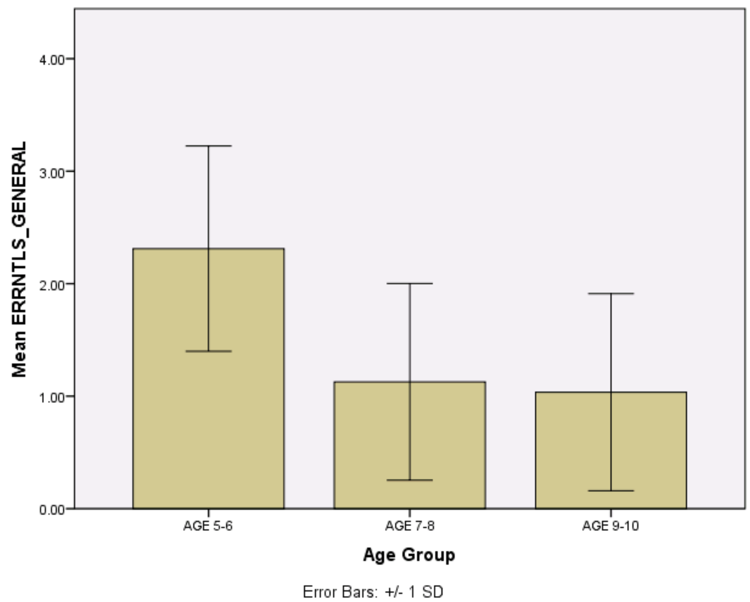


Figure 4.4.7. Average time to recover from one instance of tracking loss, for each Age Group.

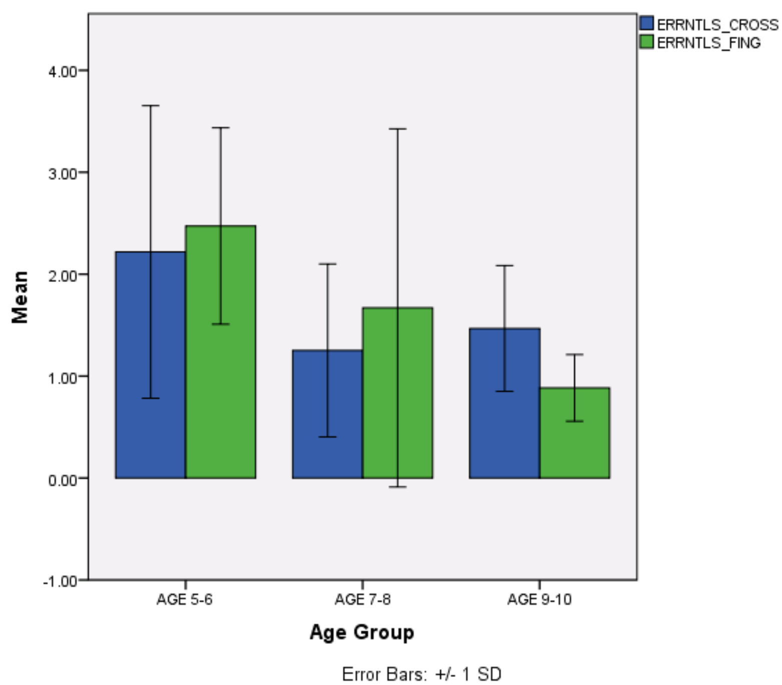


Figure 4.4.8. Average time to recover from one instance of tracking loss, for Crosshair and Finger conditions.

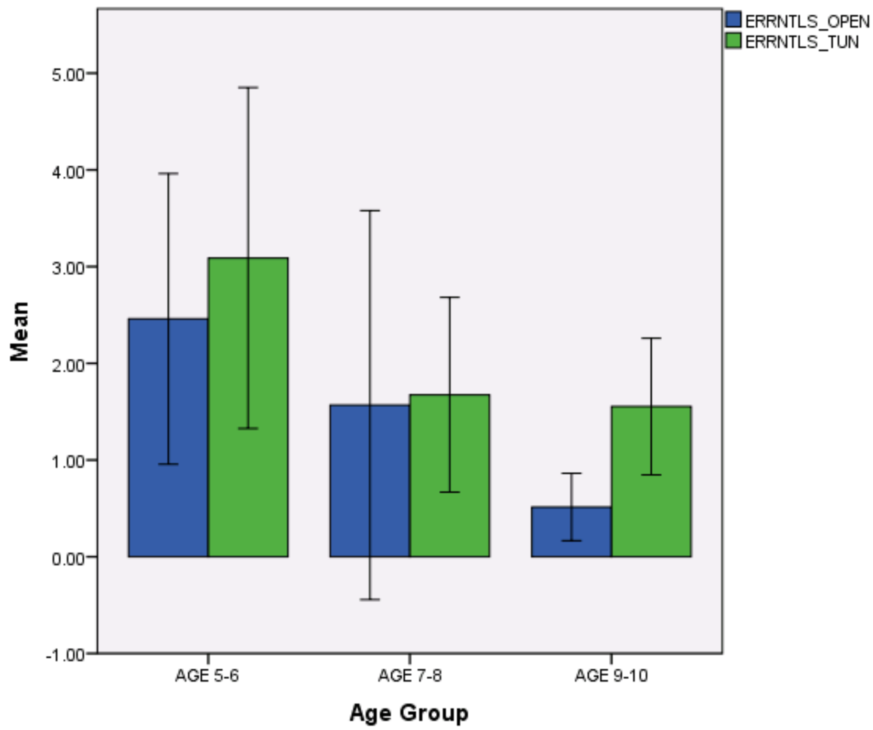


Figure 4.4.9. Average time to recover from one instance of tracking loss, for No Tunnel (open) and Tunnel conditions.

Overall this analysis indicates that 5-6 year old children recover more slowly from tracking losses than 7-8 and 9-10 year old children, but that 7-8 and 9-10 year old children are not statistically different from each other. The type of Selection Type or Movement Difficulty does not have a statistically significant effect on time to recover tracking.

For the descriptive statistics related to this metric, please see Appendix C. The data informs the research questions as follows:

RQ1-4: Does speed of AR tracking recovery differ between age groups?

Hypothesis: *Younger children will be slower at recovering the AR tracking.*

Results: *Overall, 5-6 year old children were significantly slower at recovering tracking than compared to 7-8 and 9-10 year olds; there was significant difference between 7-8 and 9-10 year olds.*

RQ2-4: Does speed of AR tracking recovery differ between interaction techniques?

***Hypothesis1**: Interaction techniques that involve independent hand movements will lead to lower tracking recovery speed.*

***Results**: No significant difference was found between interactions involving independent movements (Finger selection) vs. coordinated movements (Crosshair selection).*

***Hypothesis2**: Interaction techniques that involve whole-body movement will lead to lower tracking recovery speed.*

***Results**: No significant difference was found between interactions involving whole-body movement (Tunnel levels) vs. no whole-body movement (No Tunnel levels).*

4.4.3.4 Number of Selection Errors

When analyzing the dependent measure Number of Selection Errors, the data violated the normality assumption of the parametric test. Nonparametric analysis yielded the same significant results as the parametric analysis. The parametric analysis is reported below.

No significant main effect of Selection Type was found ($F(1,39)=0.121$, $p=0.73$); overall, the average number of errors in selecting items with the Finger was not significantly different than with the Crosshair. **A significant main effect of Movement Difficulty was found** ($F(1,39)=34.70$, $p<0.001$). Overall, the number of selection errors in Tunnel conditions ($M=1.1$, $SD=0.9$) was significantly higher than No Tunnel conditions by 49% on average. **No significant main effect of Age Group was found** ($F(2,39)=1.32$, $p=0.28$). Overall, the average number of selection errors was not significantly different across different age groups. **No significant interaction effects** were found between these variables.

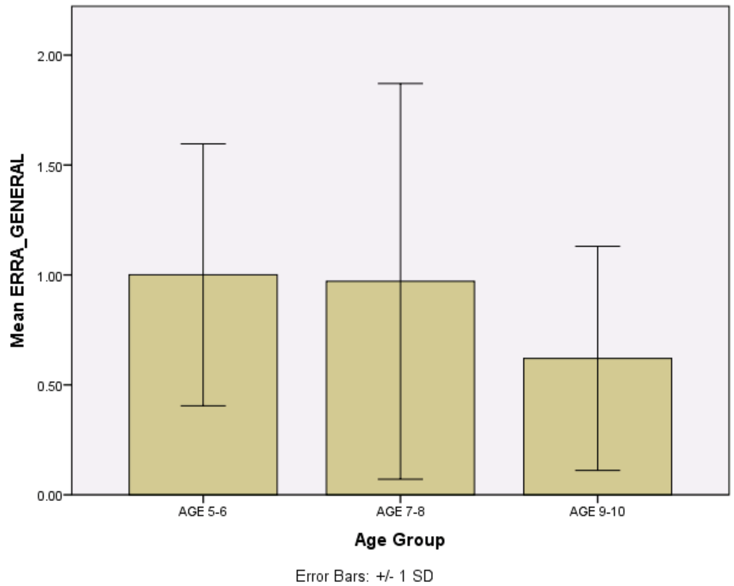


Figure 4.4.10. Average number of selection errors (per lemon), for each Age Group.

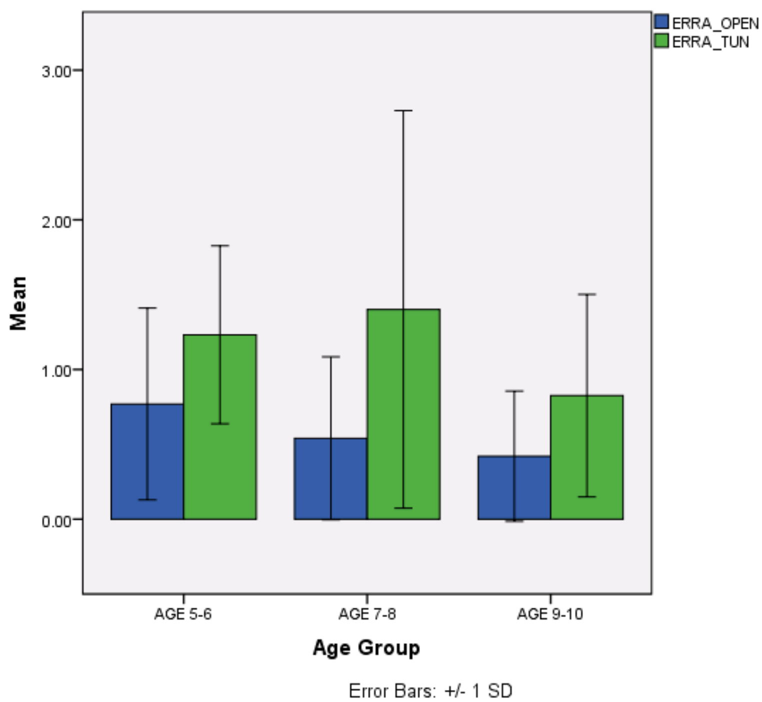


Figure 4.4.11. Average number of selection errors (per lemon), for No Tunnel (open) vs. Tunnel levels.

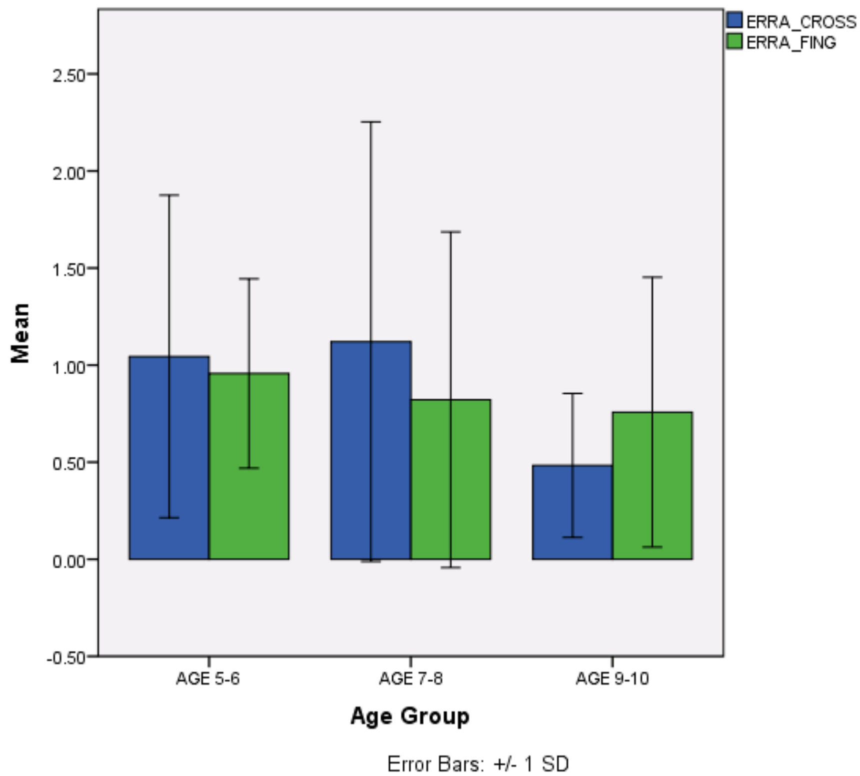


Figure 4.4.12. Average number of selection errors (per lemon), for Crosshair vs Finger conditions.

Overall this analysis indicates that the number of selection errors in Tunnel levels are significantly more than in No Tunnel levels. No other significant effects were found.

For the descriptive statistics related to this metric, please see Appendix C. The data informs the research questions as follows:

RQ1-2: Does selection accuracy differ between age groups?

Hypothesis*: Younger children will be less accurate at performing selection tasks.*

Results: *No significant differences in accuracy were found between different age groups.*

RQ2-2: Does selection accuracy differ between interaction techniques?

Hypothesis1*: Interaction techniques that involve independent hand movements will lead to lower accuracy.*

Results: *No significant difference was found between interactions involving independent movements (Finger selection) vs. coordinated movements (Crosshair selection).*

Hypothesis2: *Interaction techniques that involve whole-body movement will lead to lower accuracy.*

Results: *Overall, interaction techniques that involve whole-body movement (Tunnel levels) led to significantly higher errors than non-whole-body movement conditions (No Tunnel levels).*

4.4.4 Analysis of Physical and Cognitive Development

The analysis of physical and cognitive development investigates the variables of Spatial Relations Test, Visuomotor Precision Test, Block Construction Test, and Hand Size, and how these correlate to the metrics of AR performance (Task Completion Time, Number of Tracking Losses, Time to Recover Tracking, and Number of Selection Errors) under the four different AR interaction conditions.

I hypothesized that children who show better AR performance would show better scores on developmental tests, thus I performed one-tailed Pearson correlations to determine relationships between developmental tests and AR performance metrics. Children's age is significantly related to children's performance (Section 4.4.3), and children's experience with technology is significantly related to some performance metrics (Section **Error! Reference source not found.**), therefore I controlled for the effects of these two variables before the correlational analysis. To control for the effects of age, I used linear regression to remove the effect of age from the AR performance metrics, and I used age-standardized scores for the participants' developmental tests. To control for the effects of exposure to technology, I used linear regression to remove the effects of this variable from both metrics of developmental tests and AR performance metrics, prior to the correlation.

4.4.4.1 Spatial Relations Test

This test was significantly associated with only one experimental condition. For levels involving Crosshair selection under Tunnel conditions, the Spatial Relations Test was **significantly negative correlated to Time to Recover Tracking** (Spearman rho=-0.688, p<0.001). Children who scored higher on this developmental test were associated with lower time to recover tracking under these conditions. There were no other statistically significant correlations between this developmental test and any other AR performance metrics.

4.4.4.2 Visuomotor Precision Test

The visuomotor precision test appeared correlated with several conditions involving Crosshair selection. For levels involving Crosshair selection under No Tunnel conditions, the Visuomotor Precision Test was **significantly negatively correlated to Number of Selection Errors** (Spearman rho=-0.349, p=0.012), and significantly negatively correlated to Task Completion Time (Spearman rho=-0.280, p=0.036). For levels involving Crosshair selection under Tunnel conditions, the Visuomotor Precision Test was **significantly negatively correlated to Number of Tracking Losses** (Spearman rho=-0.312, p=0.022). In all these cases, children who scored higher values on the Visuomotor Precision Test exhibited better performance on the outlined AR performance metrics. There were no other statistically significant correlations between this developmental test and any other AR performance metrics.

4.4.4.3 Block Construction Test

For levels involving Finger selection under Tunnel conditions, the block construction test was **significantly negatively correlated to Task Completion Time** (Spearman rho=-0.422, p=0.003), and **significantly negatively correlated to Number of Selection Errors** (Spearman rho=-0.270, p=0.044).

For levels involving Crosshair selection under No Tunnel conditions, this test was **significantly negatively correlated to Time to Recover Tracking** (Spearman rho=-0.542, p=0.034). In all these cases, children who scored higher values on the developmental test exhibited better performance on the outlined AR performance metrics. There were no other statistically significant correlations between this developmental test and any other AR performance metrics.

4.4.4.4 *Developmental Test Summary*

The results from the developmental test correlations inform the research question as follows:

RQ2-5: Does child development correlate with performance under different interaction techniques?

> ***Hypothesis1****: Performance on interaction techniques that involve independent hand movements will be inversely correlated to tests of fine motor skills and physical manipulation.*
>
> ***Results****: When independent hand movement (finger selection) was required, the physical manipulation test (block construction) was inversely correlated with performance on completion time and accuracy on the whole-body movement level. When coordinated hand movement was required (crosshair selection), the fine motor skills test (visuomotor precision) was inversely correlated with performance on selection accuracy and completion time on the non whole-body movement level; and with performance on number of tracking losses on the whole-body movement level. The physical manipulation test (block construction) was also inversely correlated with coordinated hand movement performance on time to recover tracking on the whole-body movement level. The 2D spatial skills test (Spatial Relations) was significantly inversely correlated with coordinated hand movement performance on time to recover tracking on the whole body-movement level.*
>
> ***Hypothesis2****: Performance on interaction techniques that involve whole-body movement will be inversely correlated to tests of fine motor skill (visuomotor*

precision test), physical manipulation skill (block construction test), and spatial relations skill (2D spatial relations test).

***Results**: When whole-body movement was required (tunnel conditions), the physical manipulation skill was inversely correlated with performance on independent hand movement conditions (on the metrics of completion time and accuracy); the fine motor skill was inversely correlated with performance on coordinated hand movement conditions (on metrics of number of tracking losses); finally, the spatial relations skill was inversely correlated with performance on coordinated hand movement conditions (on the metric of time to recover tracking). When whole-body movement was not required, the fine motor skill was inversely correlated with performance on coordinated hand movement (on metrics of task completion time and number of errors); in the same condition, the physical manipulation skill was inversely correlated (to metric of time to recover tracking).*

4.4.5 Subjective Measures

Subjective measures of Fun, Ease of Use, and Comfort were collected using a modified Smileyometer scale, where agreement ratings ranged between Absolutely Not (1), Not Really (2), So-So (3), Yes a bit (4), and Yes very much (5).

Figure 4.4.13. Comfort question administered using Smileyometer

To analyze subjective measures, I performed 3 nonparametric tests for each metric. Kruskal-Wallis H test was used to test for effects of Age Group, and Wilcoxon signed-rank tests for the effects of Movement Difficulty and Selection Type.

4.4.5.1 Self Reported Fun

No **statistically significant differences in self-reported Fun,** across any of the factors of Age Group, Selection Type and Movement Difficulty. The average level of Fun was 4.3 / 5 (SD=0.8). Descriptive statistics indicate a decreasing trend in fun levels as children become older, possibly because the game storyline may be too simple for older children.

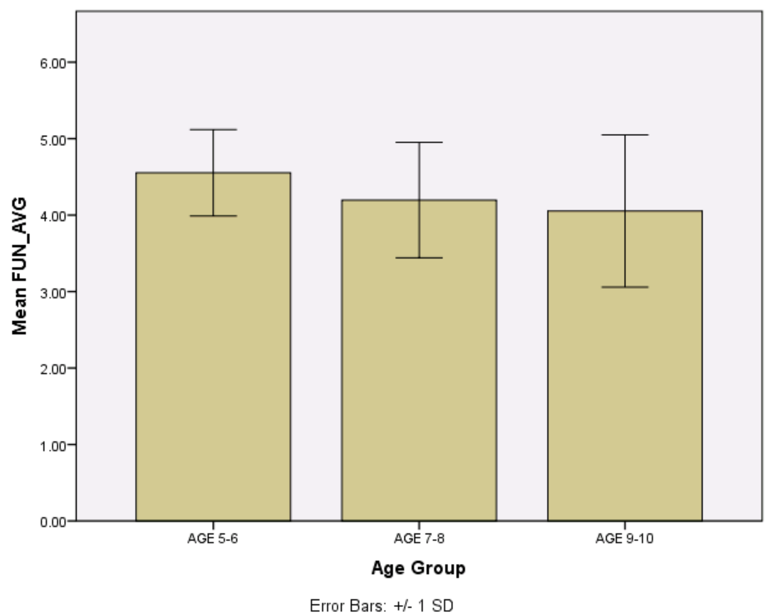


Figure 4.4.14. Average self-reported fun (/5, per level), for each Age Group.

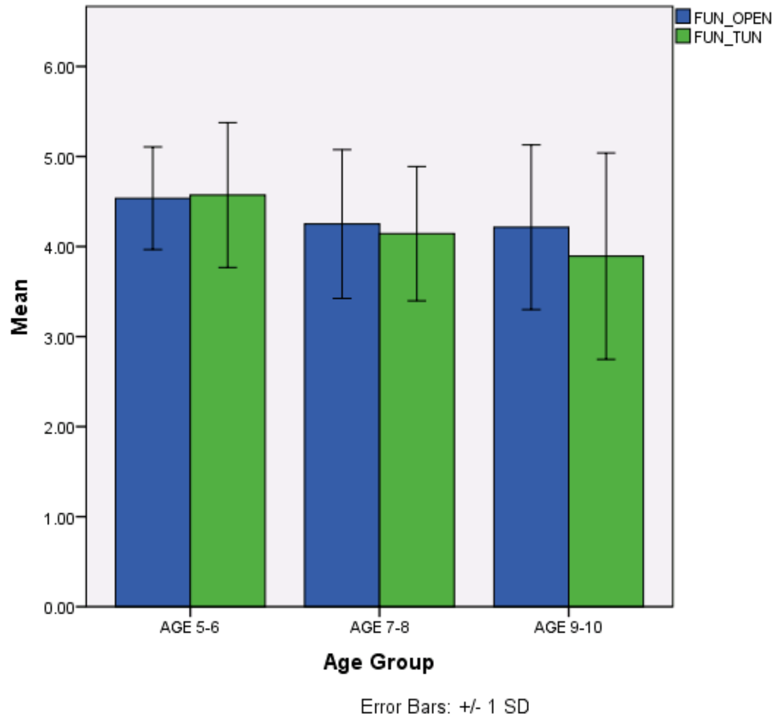


Figure 4.4.15. Average self-reported fun (/5, per level), for No Tunnels vs Tunnels levels.

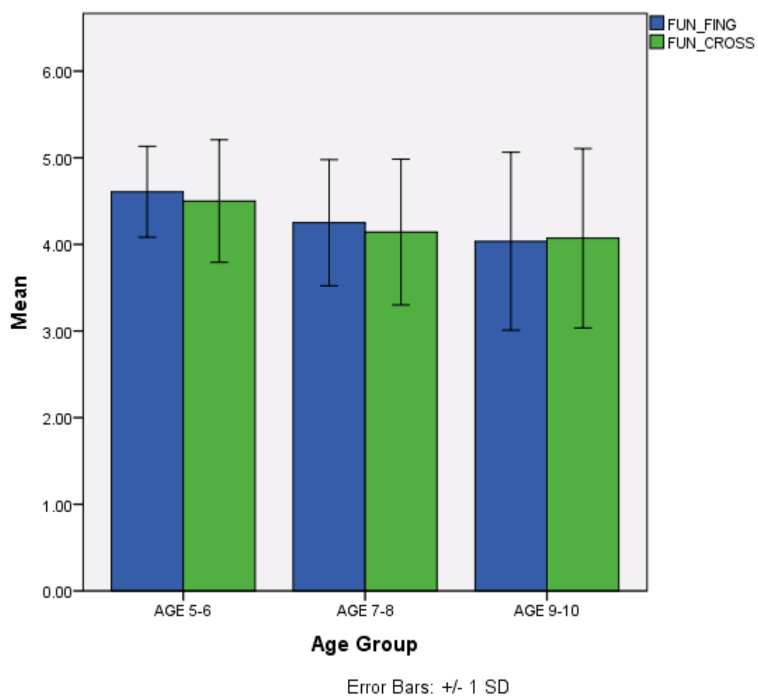


Figure 4.4.16. Average self-reported fun (/5, per level), for Finger vs Crosshair levels.

4.4.5.2 Self Reported Ease of Use

There was a statistical difference in reported ease of use between Selection Types ($z=3.366$, $p=0.001$): Finger conditions ($M=4.4$, $SD=0.6$) were rated an average 9% more easy to use than Crosshair conditions. There was also a statistical difference in reported ease of use between Movement Difficulties ($z=2.274$, $p=0.023$): No Tunnel conditions ($M=4.3$, $SD=0.7$) were rated an average 2% more easy to use than Tunnel conditions.

Descriptive statistics indicate that ease of use ratings remain relatively similar between age groups. Looking at each Selection Type, the ratings of ease of use for Finger levels appear to increase with age, while the ratings for ease of use for Crosshair levels appears to decrease.

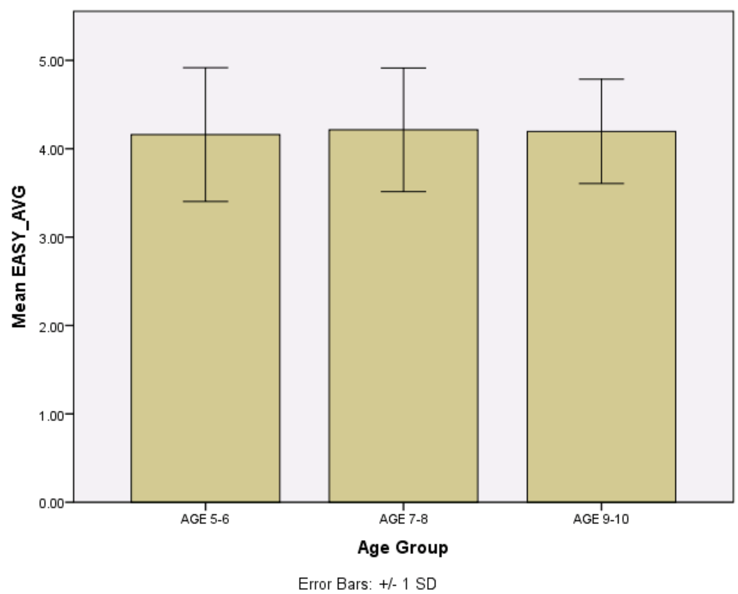


Figure 4.4.17. Average self-reported ease of use (/5, per level), for each Age Group.

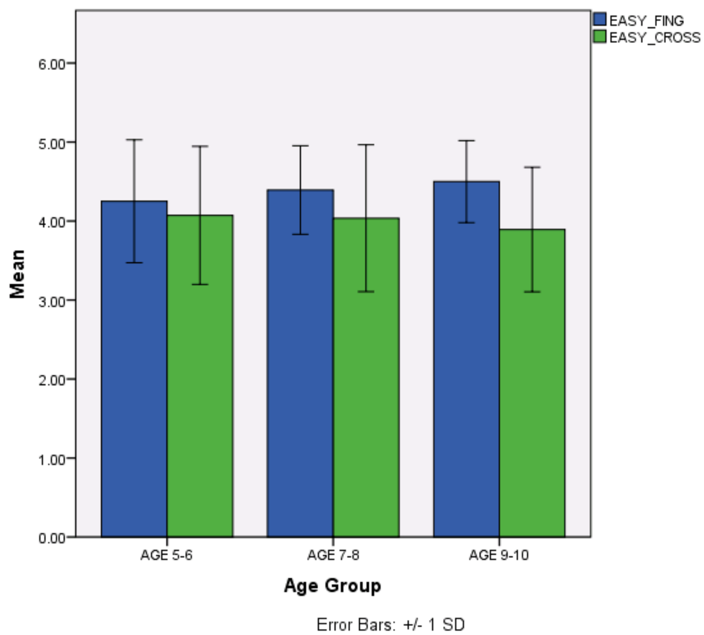


Figure 4.4.18. Average self-reported ease of use (/5, per level), for Finger vs. Crosshair levels.

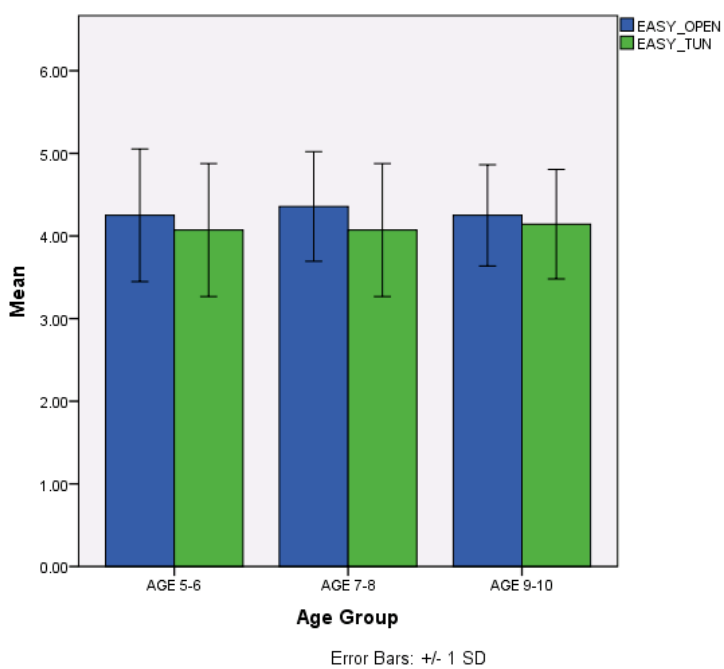


Figure 4.4.19. Average self-reported ease of use (/5, per level), for No Tunnel vs. Tunnel levels.

4.4.5.3 Self Reported Comfort

No statistically significant differences in self-reported Comfort, across any of the factors of Age Group, Selection Type and Movement Difficulty. The overall self-reported comfort was 4.1 / 5 (SD=0.8).

Descriptive statistics indicate an increasing overall trend in comfort levels as children become older. Based on descriptive statistics, Crosshair interaction appears more comfortable for 5-6 year old children; for the older age groups, Finger interaction appears more comfortable. This might be due to 5-6 year olds' difficulty of manipulating the phone with their hands in Finger selection mode. During informal observations during the study, I observed young children having trouble gripping the phone in a manner comfortable for touching the screen with a finger; while for Crosshair mode they could curl their hands around the screen, potentially causing this trend of Crosshair being more comfortable; in older children they seemed more comfortable touching the screen in Finger mode, as their hands were larger. Another explanation may be that young children's hands are more unsteady, thus they may move the screen while attempting to touch with the finger, thus feeling more discomfort.

The Tunnels condition appears to have a non-significant descriptive trend, moving from being less comfortable than No Tunnels for 5-6 year olds, to being more comfortable than No Tunnels for 9-10 year olds. During the studies I observed that 9-10 year olds hold a poor posture (e.g., back/neck bent in uncomfortable-looking positions), thus the movement in Tunnel levels may cause loosening up of the body's uncomfortable postures, leading to more comfort. Another explanation for why Tunnels may be more comfortable than No Tunnels for older children, is that older children reportedly prefer Tunnels over No Tunnel levels, thus they may report No Tunnel levels as less comfortable because they are less engaging.

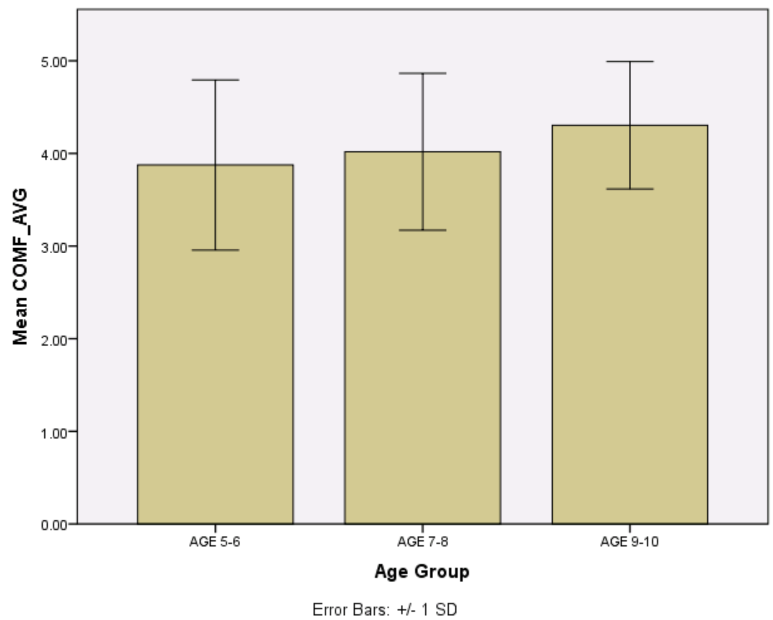


Figure 4.4.20. Average self-reported comfort (/5, per level), for each Age Group.

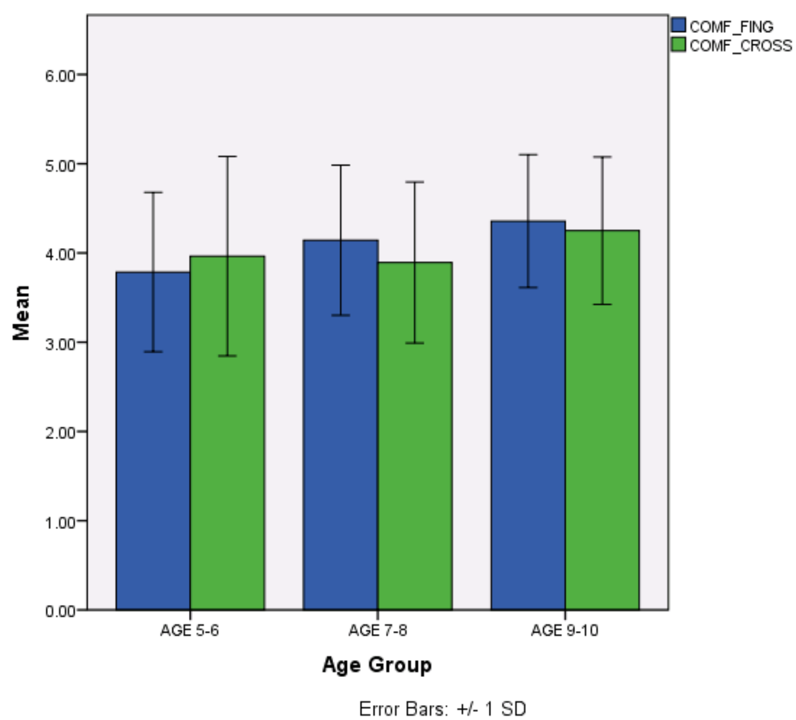


Figure 4.4.21. Average self-reported comfort (/5, per level), for Finger vs Crosshair levels.

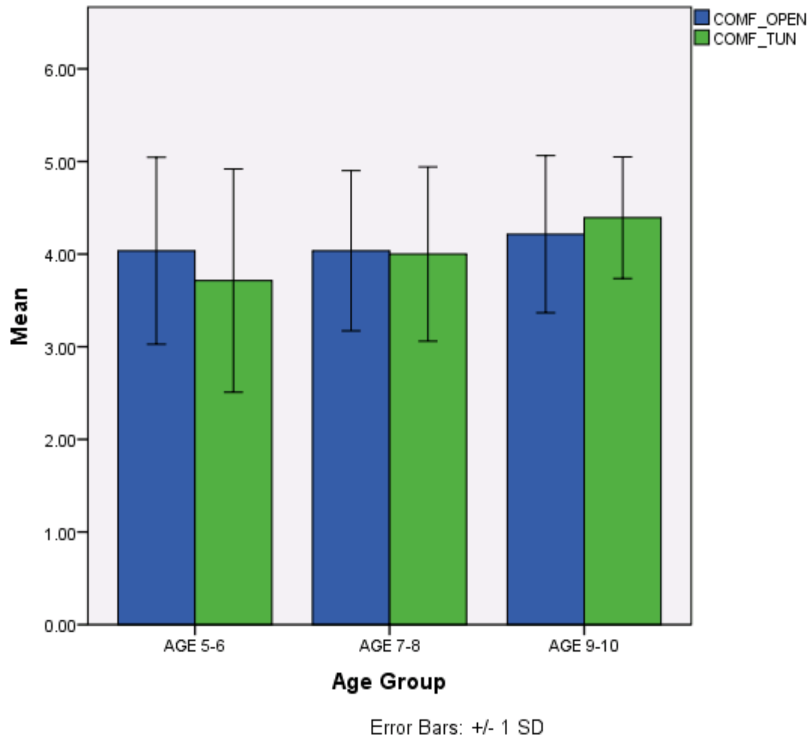


Figure 4.4.22. Average self-reported comfort (/5, per level), for No Tunnel vs Tunnel levels.

4.4.5.4 Summary of Subjective Measures

The analysis above provides answers to the research questions as follows:

RQ1-5: Does children's self-reported fun, ease-of-use, and comfort change with age?

> ***Hypothesis1****: As age increases, there will be an increase in self reported ease-of-use.*
>
> ***Results****: Ease-of-use did not significantly differ between age groups.*
>
> ***Hypothesis2****: As age increases, there will be an increase in self reported comfort.*
>
> ***Results****: Comfort did not significantly differ between age groups.*
>
> ***Hypothesis3****: It is unclear if fun would change across age groups.*
>
> ***Results****: Fun did not significantly differ between age groups.*

RQ2-6: Does children's self-reported fun, ease-of-use, and comfort change between interaction techniques?

***Hypothesis1**: Interaction techniques requiring independent hand movements or whole-body movements will yield less self-reported ease-of-use.*

***Results**: Ease of use was reported significantly higher for the interaction technique requiring independent hand movements (Finger selection). Ease of use was reported significantly higher for conditions where no whole-body-movement was required (No Tunnel levels).*

***Hypothesis2**: Interaction techniques requiring independent hand movements or whole-body movements will yield less self reported comfort.*

***Results**: Comfort did not significantly differ between interaction technique conditions.*

***Hypothesis3**: It is unclear if fun would change across interaction techniques.*

***Results**: Fun did not significantly differ between interaction technique conditions.*

4.4.6 Analysis of Observed Usability Problems

4.4.6.1 Coding Scheme Development and Validity

A coding scheme was developed prior to the main study, in order to identify problematic child behaviors based on video recordings. The codes were developed through a process of initial coding scheme creation, followed by iterative refinement by coding videos of children playing with the experimental game. Appendix B contains the final coding scheme and documentation used for training video coders.

The initial coding scheme was developed from combining the DEVAN coding scheme for young children's usability in desktop computer-based games (Barendregt and Bekker 2006), and my AR Child Usability Framework (Radu and MacIntyre 2012). The goal of the coding scheme in my research was to identify behaviors which point at usability problems when children experience handheld augmented reality games. Specific codes from the DEVAN coding scheme were selected for their ability to identify generic child behaviors indicative of problems (for example, when children appeared frustrated or bored); the AR Child Usability Framework was used to identify other behaviors specifically related to augmented reality. For example, the final coding scheme identifies discrete events such as whenever children expressed frustration, when they dropped the phone, when they lost tracking (and for what reason); and also time series events such as the type and duration of grips. The coding scheme was developed by myself working with a team of undergraduate video coders, and a large part of its refinement was based on video recordings from the pilot studies. Once the coding scheme showed stability and non-ambiguity, we applied it to coding the real experiment videos.

Inter-rater agreement validity was based on Cohen Kappa, calculated using the Quera and Bakeman GSEQ algorithm for calculating agreement for event-based agreement for time-based sequences (Quera and Bakeman 2000). The videos from all children the main study were coded by two coders (myself and one undergraduate student), overlapping on 10% of the videos. On each of the overlapping videos, both coders achieved a Cohen's Kappa of 0.80 or greater (corresponding to 97% or greater agreement). The final coding scheme is provided in Table 4.4.2 (with details about each code in Appendix B). The coding scheme contains 39 different code types in 7 families.

Table 4.4.2. The video coding scheme developed and used in the ARC study.

Category: Game Events LEVELSEC_S / E - Level Started/Ended (record on first green spell appearance in level) TL_* – Tracking Lost (>1s) – insert reason
Category: Help and Interruptions HLPME~ - Child asks for help HLPV_SINGLE~ - Experimenter gives quick verbal help INTH1_S~ /E – Gameplay interrupted: When experimenter helped for a period, just verbally INTH2_S~ /E – Gameplay interrupted: When experimenter helped for a period, and had to take away the phone, or move the paper, or touch the child IGN~ - Child ignored instructions INTC_S~ /E – Gameplay interrupted: Child does something (verbal or nonverbal) which causes their gameplay to be interrupted (e.g., looking at experimenter) SAIDU~ – Child or Experimenter said something that is Unknown (can't be understood)
Category: Mental and Emotional States SPACEV / NV~ - Child gives indication of space (verbal or non-verbal) FRUV / NV~ - Indication of frustration or dislike (verbal or non-verbal) SUGG~ - Child makes a suggestion (verbally) CONF~ - Indication of confusion (verbally) BOR~ - Indication of boredom (verbally) LIKE~ - Indication of liking (verbally) AHA~ - Indication of "aha" moment (verbally) TIRV~ - Indication of physical tiredness (verbally)
Category: Codes during Tutorial THLP_FINGER / _TUNNELS / _CROSSHAIR~ - Experimenter needs to give help about touching with finger / looking into the tunnels / using crosshair (first time only) SPACEV~ / SPACENV~ *(as above)* LEVELSEC_S / LEVELSEC_E *(as above)* SAIDU~ (as *above*) LIKE~ (as *above*) PHONEDROP_* (as *below*)
Category: General Movement SCRATCH - Scratches (code for each 2s) TIRNV_* - Tiredness, muscle strain, or stiffness observed non-verbally as: HSTR - Fingers / hand / arm stretched HSHAKE - Hand shaken BSTR - Body stretching BSIT – Body sitting down ELT - Elbow or hand is resting on table PHONEDROP_* - Phone is dropped or slips – partial or full PHONEDOWN - Puts the phone down on table BUMP~ - Bumps or trips body into physical object (or trips over themselves)
Category: Grips *Grip modifiers are: CRAB, CURL, STR, CORN, BOTTOM, X* R**GSW**_* / L**GSW**_* / **RL**GSW_*r_*l / **X**GSW_*r_*l – Grip of hand has switched
Category: Backs *Back modifiers are: STR, BENT* BSW_* / XBSW_*– Back posture has switched (only if held > 3 seconds)

In order to identify types of usability problems experienced by children, I followed a 3-step process of iteratively analyzing the event codes collected from videos of children playing with the game. In the first phase, event occurrences within each code type were categorized into clusters, where each cluster was associated with a usability problem theme and severity. At this stage, some events were ambiguous and thus were assigned to multiple categories (for example, when the experimenter instructed the child to move closer to fix the tracking loss, this may indicate a difficulty in knowing how to orient the body in relation to the gameboard, or a difficulty knowing how to recover tracking, thus the event was counted as both categories). The severity rating for each cluster was assigned according to how much help a child would require to fix the usability problem (Table 4.4.3.). In the second phase of the process, a set of overarching usability problems was identified by aggregating clusters of events - the final list of problems is summarized in Table 4.4.4. Finally, the usability problems were categorized under each of the four developmental areas of the AR Child Usability Framework. The problems will be described along with their analysis in the following section.

Table 4.4.3. Severity ratings for the identified usability issues.

Problem Severity
0 = Very Low The issue had no visible impact on the child's gameplay.
1 = Low The issue frustrated the child, but no other gameplay impact.
2 = Medium The issue was fixed after verbal help from the experimenter.
3 = High The experimenter had to interrupt the gameplay in order to fix the issue.

Table 4.4.4. Observed usability issues, and statistically significant positive correlations (C+), negative correlations (C-), differences between tunnel and no-tunnel conditions, and other group differences (X).

ISSUE CATEGORY	NUM. CHILDREN WHO ENCOUNTERED THIS ISSUE			STATISTICALLY SIGNIFICANT EFFECTS DETECTED (P<0.05)		
	5-6 years old	7-8 years old	9-10 years old	Age	Movement: tunnels (T) vs. no tunnels (NT)	Grips
Overall				C-		
Manipulation						
Losing tracking while walking	14 (100%)	7 (50%)	6 (50%)	X	T > NT	X
Losing tracking by covering the camera with the finger	10 (71%)	11 (79%)	8 (66%)			
Strained grip	0 (0%)	5 (35%)	2 (16%)			
Dropping the phone	2 (14%)	1 (7%)	2 (17%)			
Strained body posture	0 (0%)	4 (29%)	5 (41%)	C+	T > NT	
Space						
Losing tracking by aiming the camera away from the gameboard	8 (57%)	4 (29%)	1 (8%)	C-		
Losing tracking by aiming the camera too close to the gameboard	2 (14%)	2 (14%)	0 (0%)			
Difficulty orienting body in relation to the gameboard	9 (64%)	7 (50%)	0 (0%)	X	T > NT	
Abstract thinking						
Needing initial instruction on how to use crosshair	10 (71%)	11 (79%)	7 (58%)			
Needing in-game instruction on how to use crosshair	2 (14%)	1 (7%)	0 (0%)			
Not understanding the game storyline	1 (7%)	3 (21%)	1 (8%)			
Not understanding general game mechanics	3 (21%)	2 (14%)	0 (0%)			
Difficulties interpreting tracking loss and recovering tracking	7 (50%)	2 (14%)	2 (16%)	C-	T > NT	
Attention						
Bumping or tripping	5 (36%)	5 (36%)	5 (41%)		T > NT	
Interruption due to self-distraction	3 (21%)	1 (7%)	1 (8%)			
Interruption due to scratching	3 (21%)	8 (57%)	8 (67%)	C+		X

4.4.6.2 Usability Problem Results

4.4.6.2.1 Summary of Analysis and Results

The usability issues that were identified through the qualitative analysis are presented in Table 4.4.4. and further descriptive statistics are listed in Appendix C. These issues will be discussed in turn in the following sections.

Each issue was analyzed statistically to determine significant effects of Age, Movement Difficulty, and Selection Type. I performed parametric tests of Pearson correlations and analysis of variance. However, the parametric assumption of normality was violated for all usability problems, thus I also performed nonparametric Spearman correlations, Kruskal-Wallis H tests (for the between-subjects factors) or Wilcoxon signed-rank tests (for the within-subjects factors), with Bonferroni-corrected Type I thresholds. Whenever the parametric assumptions were violated but nonparametric tests showed the same significant differences between groups, I report the parametric test results.

These issues will be discussed in the following sections. In summary, the analysis of usability issues informs the research questions as follows:

RQ1-6: Do usability issues differ between age groups?

> ***Hypothesis****: Younger children will experience higher number of usability issues.*
>
> ***Results****: There is a significant correlation indicating that the number of usability issues decreases with children's age. Such a significant correlation occurs in multiple issues related to tracking loss and body orientation.*

RQ2-7: Do usability issues differ between interaction techniques?

> ***Hypothesis****: Interaction techniques that involve independent hand movements will lead to more usability issues.*
>
> ***Results****: No significant differences were found between Finger vs Crosshair selection techniques; however, descriptive statistics show some trends.*
>
> ***Hypothesis****: Interaction techniques that involve movement difficulty will lead to more usability issues.*

Results*: Significant differences were found in several issues related to tracking loss and body movement, whereby more issues occurred in Tunnel than No Tunnel levels.*

RQ3: What types of usability issues are experienced by children in handheld-AR?

Hypothesis*: Children will experience problems that can be linked to developing areas of physical and cognitive skills*

Results*: The types of usability issues encountered by children are shown in Table 4.4.4.*

4.4.6.2.2 Relationship between Age and Overall Problem Counts

Children in all age groups experienced problems in all areas of the AR Usability Framework.

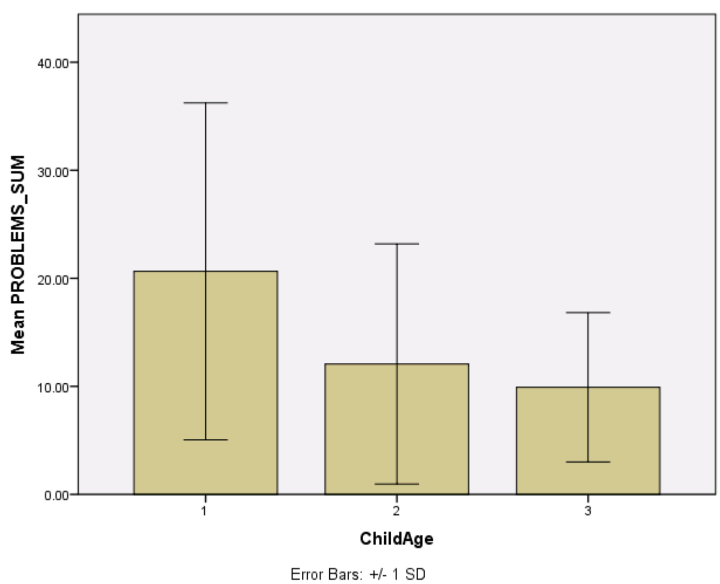


Figure 4.4.23. Average number of problems (per child) encountered in each Age Group.

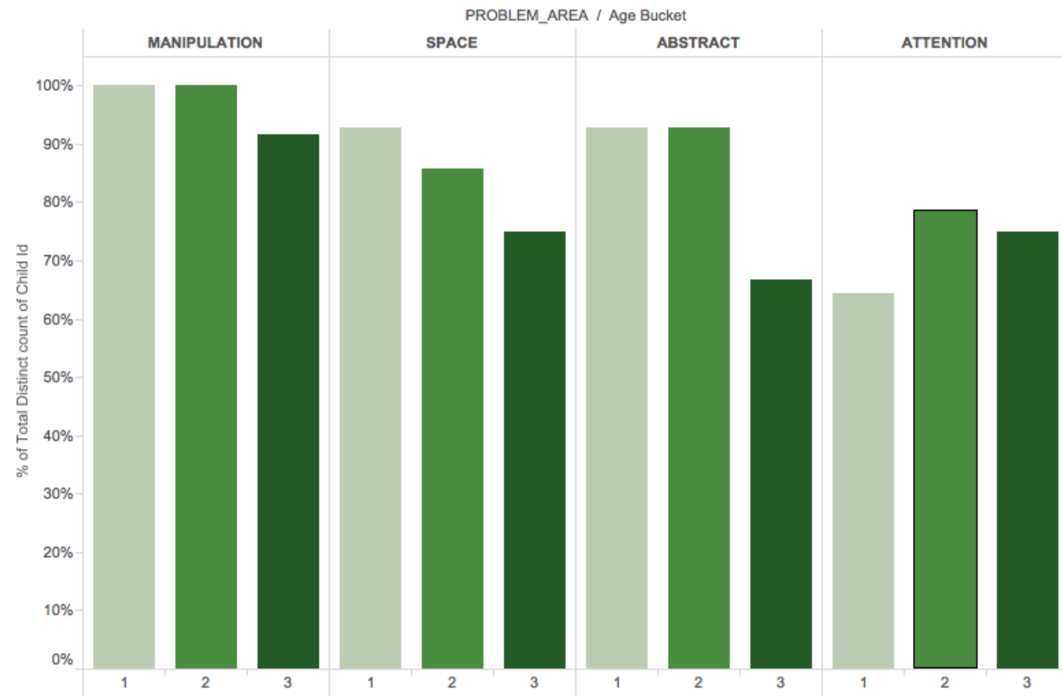


Figure 4.4.24. Percentage of children in each age group, experiencing problems in each category.

Descriptive statistics indicate young children (aged 5-6) encountered an average of 20.6 issues during the whole gameplay; children in the middle group (aged 7-8) encountered an average 41% less issues, while older children (aged 9-10) encountered an average 52% less issues compared to the 5-6 year olds. To test the hypothesis that children's age is related to the overall number of problems encountered, I performed two-tailed Spearman correlations between children's age (in months) and the overall number of errors encountered. The correlational analysis indicates that as age increases, there is a **significant inverse linear correlation between age and the overall number of errors encountered** ($R_s(40)=-0.342$, $p=0.031$). A follow-up Kruskal-Wallis H test found **no statistically significant main effect of between-subjects factor Age Group** ($p=0.094$).

When analyzing problems by their severity levels in relation to the effects of age, two-tailed Spearman correlations indicate that as age increases, there are **significant inverse linear correlations with the number of errors encountered in Severity 1** ($R_s=-0.456$, $p=0.003$), **Severity 2** ($R_s=-0.380$, $p=0.015$), **and Severity 3** ($R_s=-0.346$, $p=0.029$). This indicates that as children's age increases, there is a decrease in number of problems encountered in each of these severity levels.

4.4.6.2.3 Issues Related to Manipulation

Losing tracking while walking: Children within each age group were observed losing tracking while they walked around the gameboard. It was not possible to determine the exact cause of this event, but the event could have occurred either because the children aimed the phone away while moving, or because they put their finger in the way of the camera, or because the phone was moved too fast around the gameboard. All occurrences of this issue were rated as severity 1 (low severity); however, sometimes losing tracking would cause children to require help fixing the problem – in such cases, a higher severity event was recorded as part of another issue (either related to difficulties recovering tracking, or difficulty orienting the body, or both).

Table 4.4.5. Distribution of occurrences of losing tracking while walking.

	5-6 years old	7-8 years old	9-10 years old
Number of children experiencing this issue (all at Severity 1)	14 (100%)	7 (50%)	6 (50%)
Number of occurrences per affected child Min – Max	1-11	1-9	2-8
Number of occurrences per affected child Median	5	2	2.5

There was a significant effect of Age Group measured via Kruskal-Walis H test ($p=0.002$); post-hoc tests show that the 5-6 year olds (mean 5.2 tracking losses due to talking per child) encountered 65% more tracking losses due to walking than 7-8 year olds ($p=0.004$) and 67% more than 9-10 year olds ($p=0.01$), however there was no statistically significant difference between 7-8 and 9-10 year olds. **There was also a significant effect of Movement Difficulty** measured through Wilcoxon signed-rank test ($p<0.001$). The number of tracking losses due to walking in Tunnel levels (mean 2.8 per child) was 97% higher than in Non Tunnel levels.

Non-significant descriptive results: The descriptive statistics show 7-8 and 9-10 year old groups encountering relatively similar number of issues. Descriptive statistics show that for the 5-6 year old group, Finger selection yielded less issues on average than Crosshair selection, although there is much variation between children; for older children, the number of issues appear similar between selection groups.

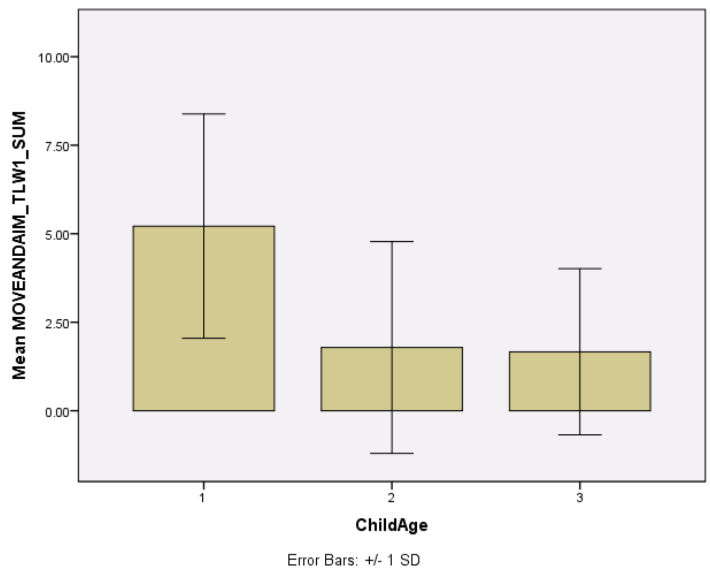


Figure 4.4.25. Total number of tracking losses due to walking, per each Age Group.

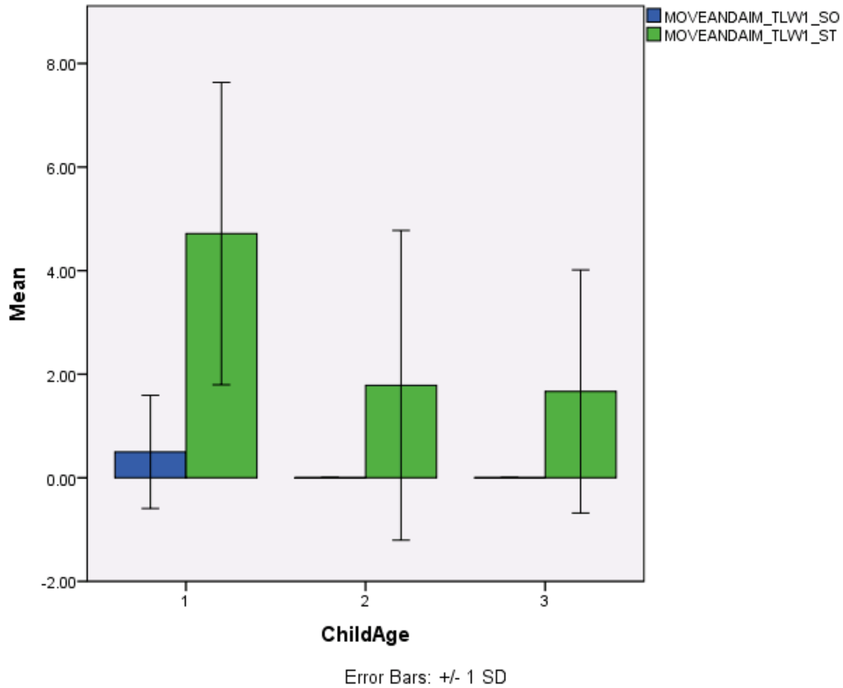


Figure 4.4.26. Total number of tracking losses due to walking, per child, for No Tunnels (blue) vs Tunnels (green).

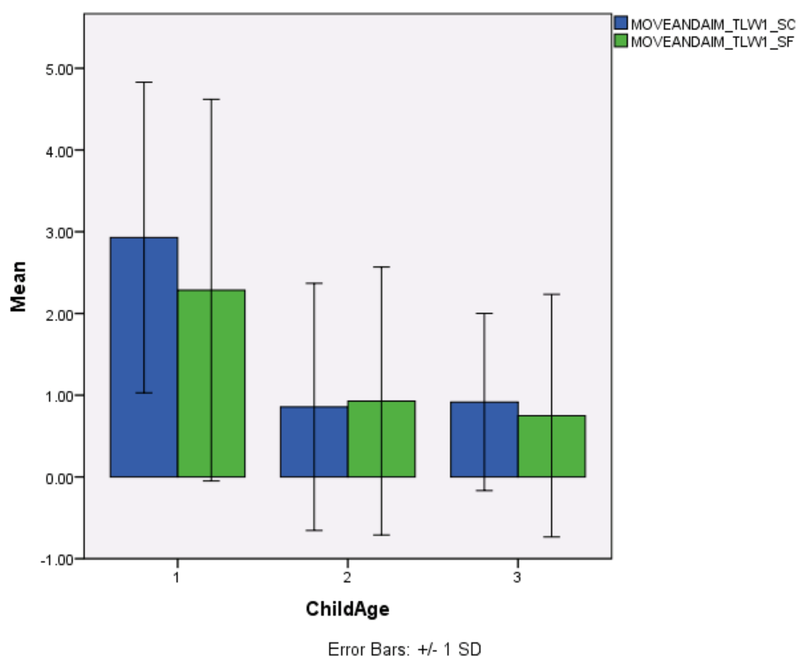


Figure 4.4.27. Total number of tracking losses due to walking, per child, for Crosshair (blue) vs. Finger (green).

Losing tracking by covering the camera with the finger: Children within each age group were observed losing tracking because they put their finger in the way of the camera. The camera was on the top left of the Atrix HD experimental phone when held in landscape mode, and covering the camera with the finger would occur if children hold the phone in a grip that encouraged fingers to be curled around the camera; furthermore, covering the camera involves some degree of inattention, whereby children would not notice when the game background would become occluded by the finger. All occurrences of this issue were rated as severity 1 (low severity); however, sometimes losing tracking would cause children to require help fixing the problem – in such cases, a higher severity event was recorded as part of another issue (either related to difficulties recovering tracking, or difficulty orienting the body, or both).

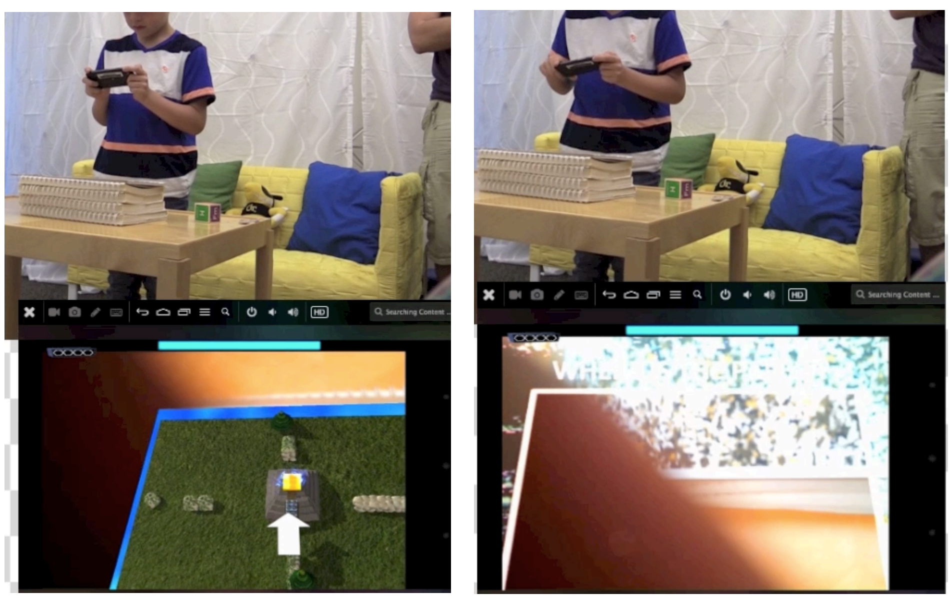


Figure 4.4.28.Child losing game tracking as they cover the camera with their finger (left), soon causing the game to lose tracking (right).

Table 4.4.6. Distribution occurrences of children losing tracking by covering the camera with the finger.

	5-6 years old	7-8 years old	9-10 years old
Number of children experiencing this issue (all at Severity 1)	10 (71%)	11 (79%)	8 (66%)
Number of occurrences per affected child Min – Max	1-18	1-11	1-5
Number of occurrences per affected child Median	3	2	1

When analyzing all children, no statistically significant correlation was found after analyzing the effects of Age, Movement Difficulty, or Selection Type.

A Spearman correlation test detected **a significant positive correlation between losing tracking with finger and the use of Straight grip in the left hand**, after the effect of Age was removed from both factors via linear regression (Rs=0.467, p=0.002).

Non-significant descriptive results: The descriptive statistics show that as children's age increases, there is a trend in decreasing the overall number of tracking loss errors due to finger occlusion. This decrease is equally visible in both movement conditions. This decreasing trend is also visible in the Crosshair conditions between all age groups, but for Finger selection the descriptive statistics show 7-8 year olds experiencing more occlusion in levels involving Finger selection than in Crosshair selection.

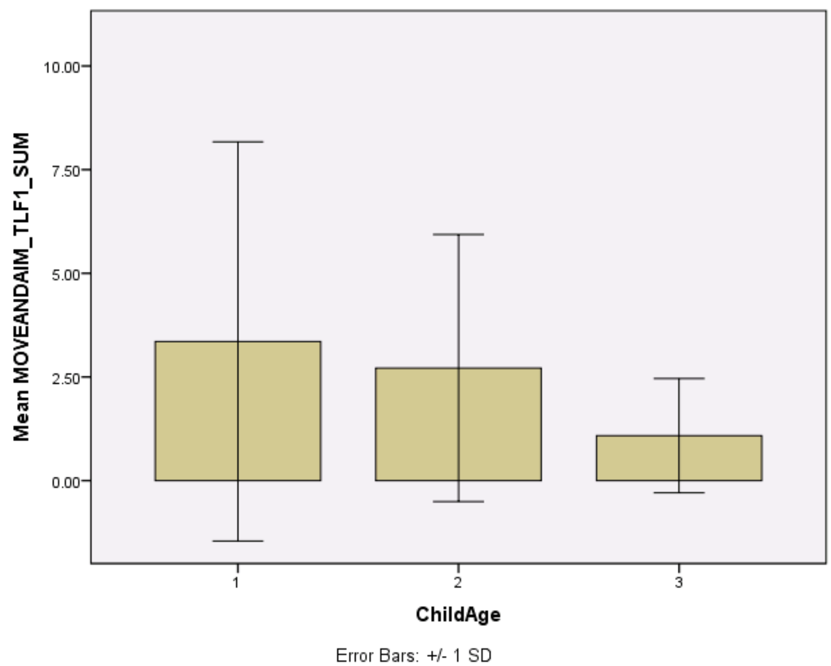


Figure 4.4.29. Average number of tracking losses due to finger occlusion, per child, for each age group.

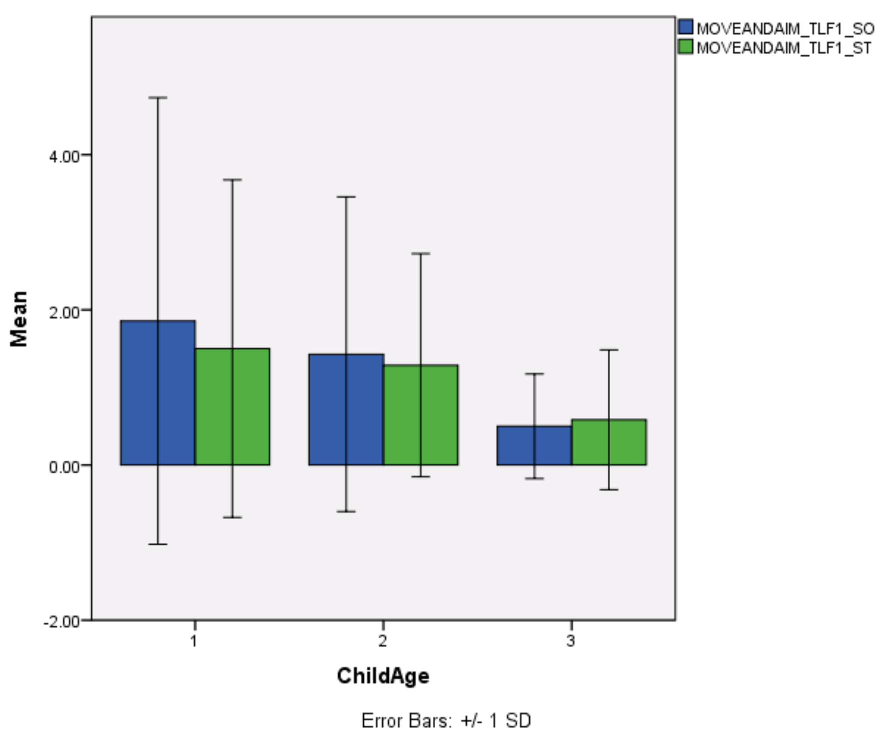


Figure 4.4.30. Total number of tracking losses due finger occlusion, per child, for No Tunnels (blue) vs Tunnels levels (green).

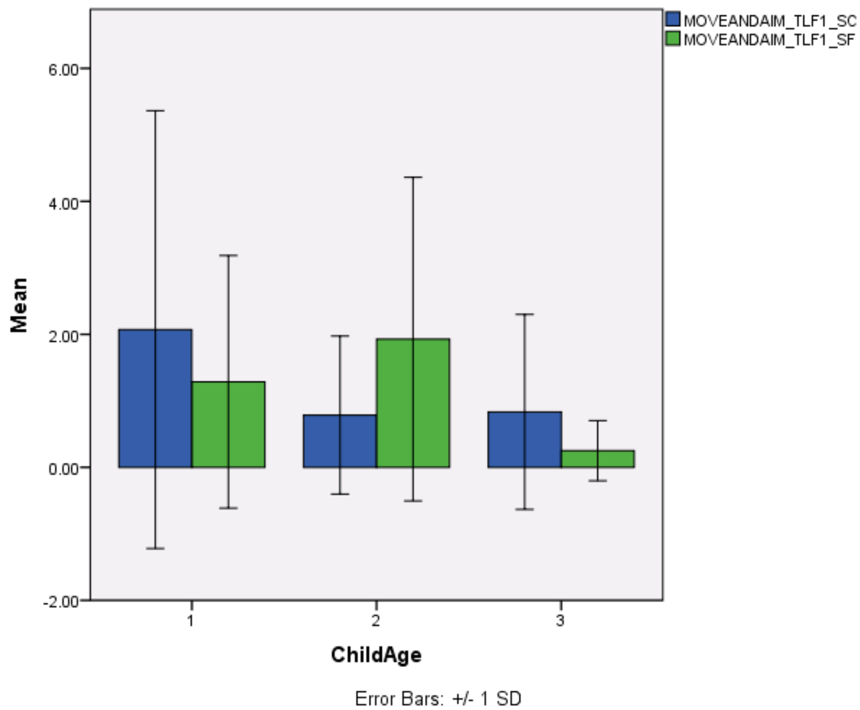


Figure 4.4.31. Total number of tracking losses due to finger occlusion, per child, for Crosshair (blue) vs. Finger (green).

Strained Grip: Children aged 7-8 and 9-10 years old were observed performing behaviors with their hands which might be indicative that their grip is being strained – this event was coded when children were observed stretching their hand, shaking their hand, or suddenly shifting their grip in a way that appeared as if they are trying to find a more comfortable grip.

Table 4.4.7. Distribution of occurrences of children showing strained grip.

	5-6 years old	7-8 years old	9-10 years old
Number of children experiencing this issue (all at Severity 1)	0 (0%)	5 (35%)	2 (16%)
Number of occurrences per affected child Min – Max	NA	1-3	2-5
Number of occurrences per affected child Median	NA	1	3.5

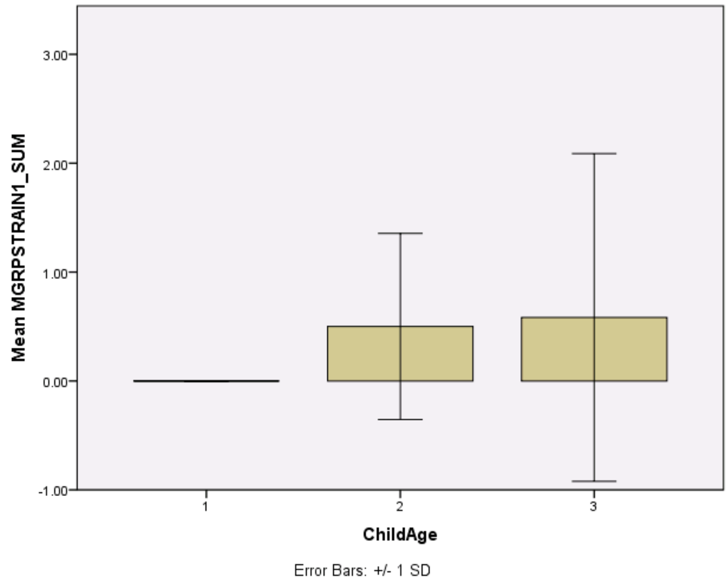


Figure 4.4.32. Total occurrences of strained grip, per child, across age groups.

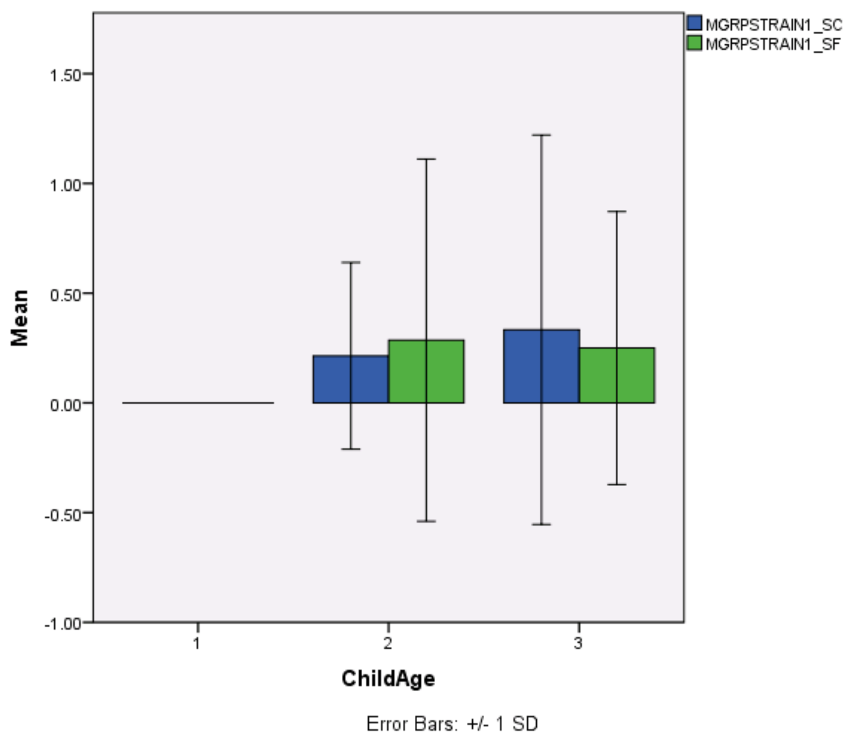


Figure 4.4.33. Total occurrences of strained grip, per child, between Crosshair selection levels (blue) and Finger selection levels (green)

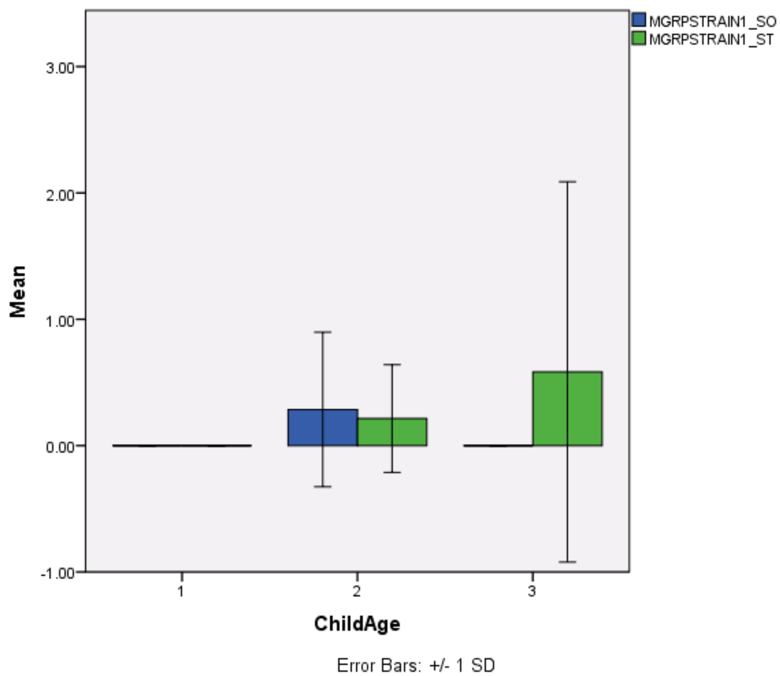


Figure 4.4.34. Total occurrences of strained grip, per child, across No Tunnel levels (blue) and Tunnel levels (green)

When analyzing all children, no statistically significant effects were found for Age, Movement Difficulty, or Selection Type.

Non-significant descriptive results: The descriptive statistics indicate that as children's age increases, there is an increase in the number of observed mean number of grip strains. The number of observed grip strains appears similar between Finger and Crosshair levels. For the children that experienced grip strains, in the 7-8 year olds, the grip strains appear similar between Tunnel and No Tunnel conditions; while in 9-10 year olds the number of grip strains appears higher and only in Tunnel conditions.

Dropping the phone: Children within each age group were observed dropping the phone either fully (the phone fell from the child's hands completely) or partially (the child caught the phone before fully falling). Informal observations noted that phone dropping occurred as children were changing or relaxing their grip.

Table 4.4.8. Distribution of occurrences of dropping the phone.

	5-6 years old	7-8 years old	9-10 years old
Number of children experiencing this issue (all at Severity 3)	2 (14%)	1 (7%)	2 (17%)
Number of occurrences per affected child Min – Max	1-1	1-1	1-1
Number of occurrences per affected child Median	1	1	1

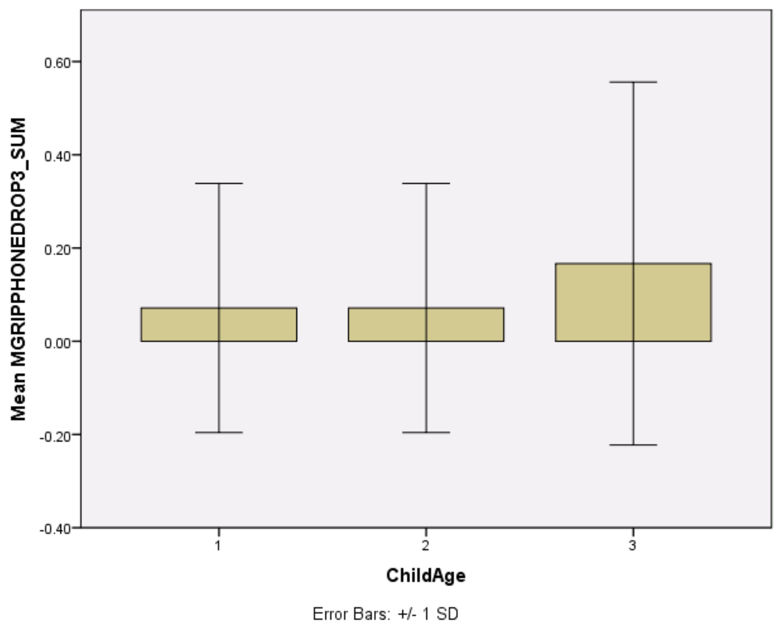


Figure 4.4.35. Total occurrences of dropping the phone, per child, across age groups.

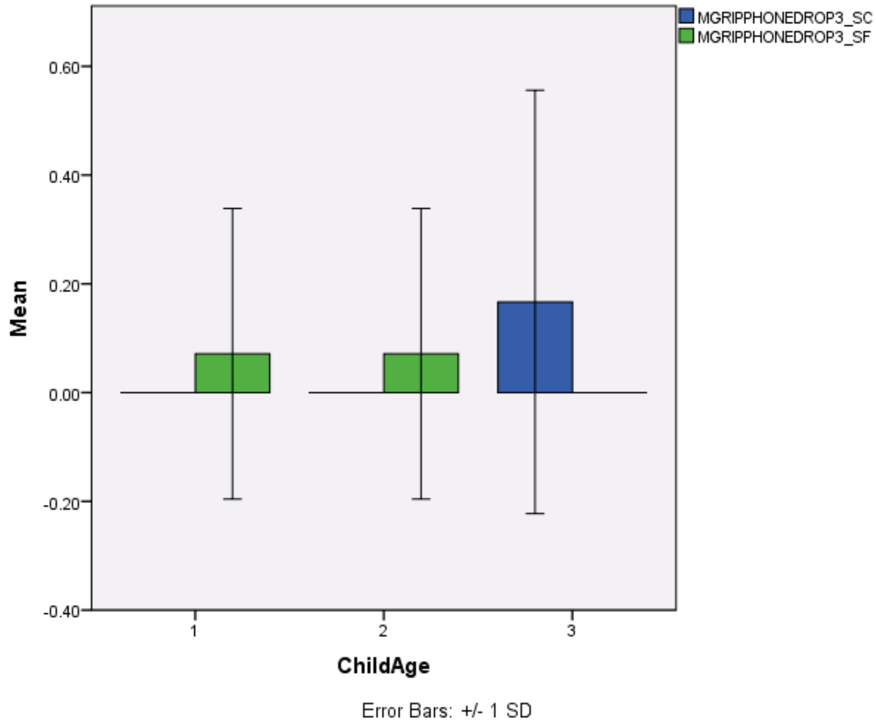


Figure 4.4.36. Total occurrences of dropping the phone, per child, between Crosshair selection levels (blue) and Finger selection levels (green)

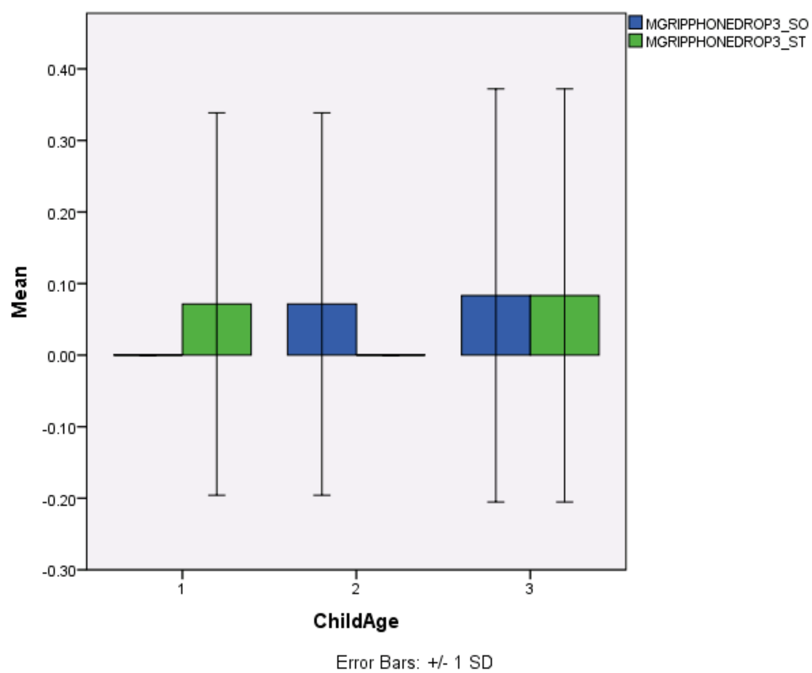


Figure 4.4.37. Total occurrences of dropping the phone, per child, across No Tunnel levels (blue) and Tunnel levels (green)

When analyzing all children, no statistically significant effects were found for Age, Movement Difficulty, or Selection Type. These results are expected due to the low number of data points.

Non-significant descriptive results: The descriptive statistics show that the number of phone drops occurs in all age groups, with higher levels in 9-10 children. When the phone was dropped in younger children, this occurred in Finger levels, involving either Tunnels or No Tunnels; in older children this occurred only in Crosshair levels involving both Tunnels and No Tunnels.

Strained Body Posture: Children aged 7-8 and 9-10 years old were observed performing behaviors with their body which might be indicative that their body is being strained – for example children bending their backs or necks in ways that appeared strained, or stretching their neck or back. The observed instances coded as Severity 0 are observations of back bending, while the observed instances coded as Severity 1 are observed stretching or straining.

Figure 4.4.38 Examples of children in what appears to be a strained body posture.

A Spearman rank correlation found **a significant positive correlation between Age and instances of strained body posture** (Rs=0.414, p=0.008), explainable by the fact that younger children played the game by walking around the gameboard, while some older children preferred to stay still while bending their back and neck more often. An analysis of variance indicated **statistically significant main effects of Movement Difficulty** (F(1,37)=4.65, p=0.038); **no other**

statistically significant main effects or interaction effects were found. Game levels involving Tunnels generated a higher number of strained body postures (mean 0.3 events per child) than No Tunnels levels (mean 0.1 events per child). This is explainable by the fact that children had to move their body more during the levels involving tunnels.

Table 4.4.9. Distribution of occurrences of strained body posture.

	5-6 years old	7-8 years old	9-10 years old
Number of children experiencing this issue	0 (0%)	4 (29%)	5 (41%)
Number of occurrences per affected child Min – Max	NA	1-7	1-14
Number of occurrences per affected child Median	NA	3	2
Children experiencing this event at Severity 0 (occurrence of back bending)	0 (0%)	3 (21%)	4 (29%)
Children experiencing this event at Severity 1 (occurrence of visible strain)	0 (0%)	3 (21%)	3 (21%)

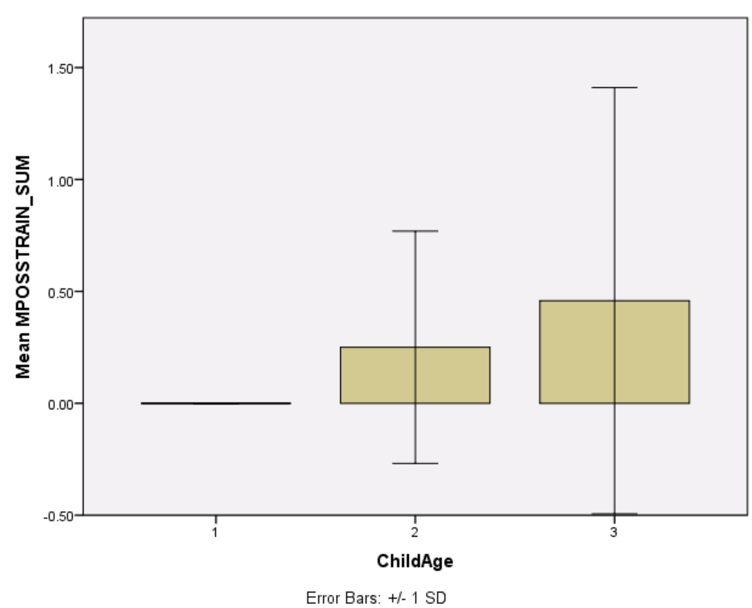


Figure 4.4.39. Total occurrences of strained body posture, per child, across age groups.

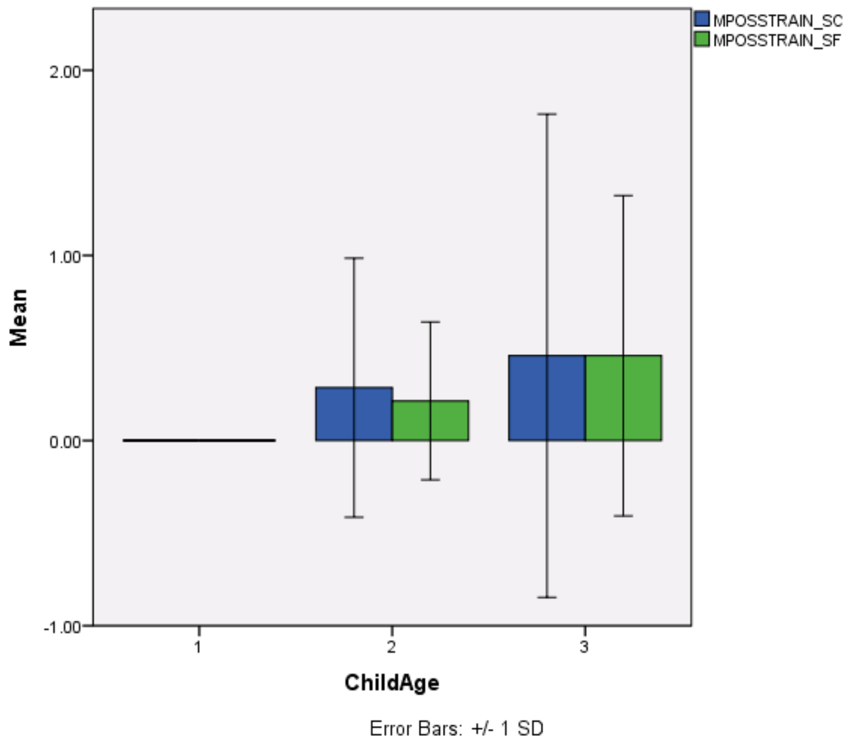


Figure 4.4.40. Total occurrences of strained body posture, per child, between Crosshair selection levels (blue) and Finger selection levels (green)

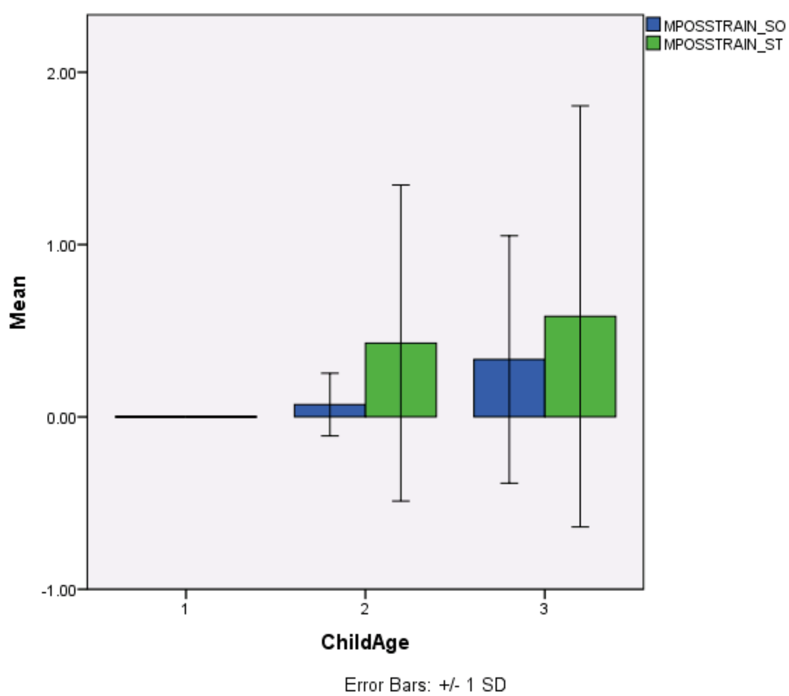


Figure 4.4.41. Total occurrences of strained body posture, per child, across No Tunnel levels (blue) and Tunnel levels (green)

Non-significant descriptive results: It is worth noting that there is an increase in observed instances of this event at Severity 1 (observed stretching or straining) in older children. There does not appear to be a difference between Finger and Crosshair levels.

4.4.6.2.4 Issues Related to Space

Losing tracking by aiming the camera away from the gameboard: Children within each age group were observed losing tracking because they aimed the camera away from the gameboard. The augmented reality game was displayed on top of the gameboard paper, and it only worked if the paper was at least partially visible to the phone camera - this due to the Vuforia augmented reality tracking technology which tracks features on the paper. During the tutorial level, children were instructed that the game works when the phone camera is pointed at the gameboard, and that if the camera is aimed away, the game would stop working until the phone is aimed back at the paper. All occurrences of this issue were rated as severity 1 (low severity); however, sometimes losing tracking would cause children to require help fixing the problem – in such cases, a higher severity event was recorded as part of another issue (either related to difficulties recovering tracking, or difficulty orienting the body, or both).

Table 4.4.10. Distribution of occurrences of losing tracking by aiming away from the gameboard.

	5-6 years old	7-8 years old	9-10 years old
Number of children experiencing this issue (all at Severity 1)	8 (57%)	4 (29%)	1 (8%)
Number of occurrences per affected child Min – Max	1-12	1-3	1-1
Number of occurrences per affected child Median	1.5	1.5	1

A Spearman rank correlation found a significant negative correlation between Age and instances of strained body posture (Rs=-0.488, p=0.001), with older children experiencing less such events. No statistical effects were found for Selection Type or Movement Difficulty.

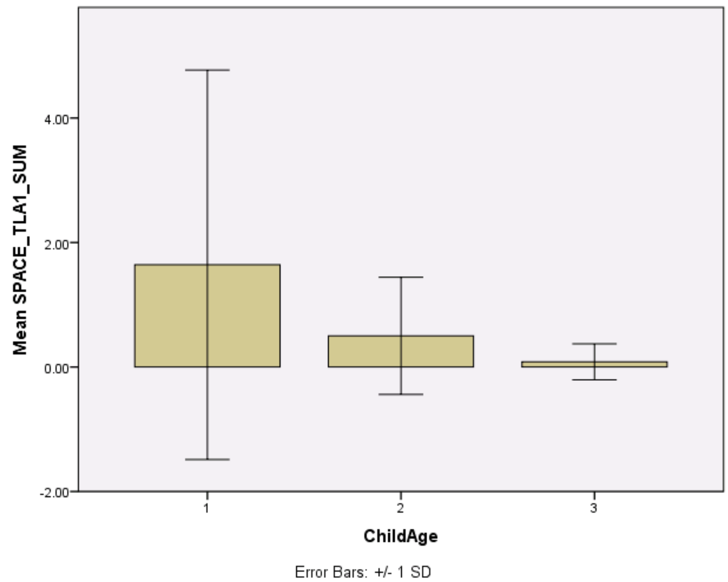


Figure 4.4.42. Total occurrences of losing tracking by aiming away, per child, across age groups.

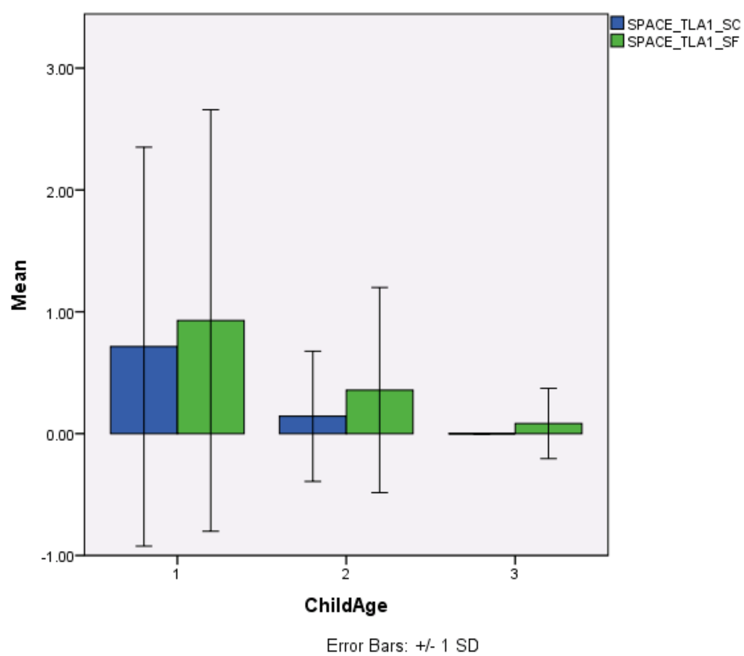


Figure 4.4.43. Total occurrences of losing tracking by aiming away, per child, between Crosshair selection levels (blue) and Finger selection levels (green)

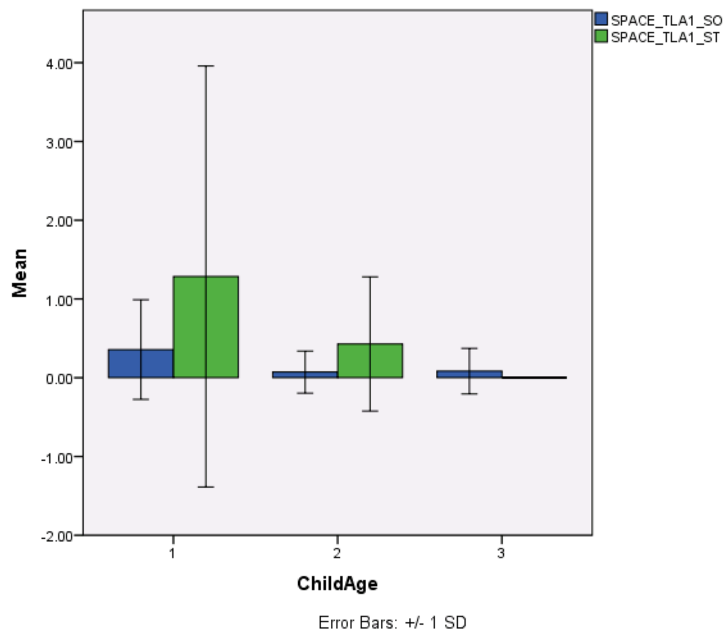


Figure 4.4.44. Total occurrences of losing tracking by aiming away, per child, across No Tunnel levels (blue) and Tunnel levels (green)

Non-significant descriptive results: For the Finger conditions there appears to be a consistently higher number of such tracking errors than Crosshair conditions. These errors also appear much more pronounced for Tunnel levels vs No Tunnel levels.

Losing tracking by aiming the camera too close to the gameboard: During the tutorial section of the experiment, children were shown that the phone will stop working if it is too close to the paper. The Vuforia augmented reality tracking technology does not work if the phone is too close to the paper, since the camera becomes out of focus and the software is unable to detect the features printed on the paper. With the experimental smartphone, Atrix HD, the tracking stops when the phone is closer than roughly 2’’ from the paper. All occurrences of this issue were rated as severity 1 (low severity); however, sometimes losing tracking would cause children to require help fixing the problem – in such cases, a higher severity event was recorded as part of another issue (either related to difficulties recovering tracking, or difficulty orienting the body, or both).

Table 4.4.11. Distribution of occurrences of losing tracking by aiming too close to the gameboard.

	5-6 years old	7-8 years old	9-10 years old
Number of children experiencing this issue (all at Severity 1)	2 (14%)	2 (14%)	0 (0%)
Number of occurrences per affected child Min – Max	1-3	1-1	NA
Number of occurrences per affected child Median	2	1	NA

Only children in the 5-6 year old and 7-8 year old age groups were observed losing tracking by this method. When analyzing all children, **no statistically significant effects were found for Age, Movement Difficulty, or Selection Type**.

Non-significant descriptive results: The descriptive statistics indicate that this issue did not occur frequently, and that the frequency of this issue decreases with age. For 5-6 year olds, this occurred in 2 children (for an average of 2 times per child); for 7-8 year olds, this occurred in 2 children (in both cases, 1 times per child); for 9-10 year olds, this issue did not occur in any children. The issue does appear to be more frequent in Tunnel levels (occurring a total of 5 times) than in No Tunnel levels (occurring only once).

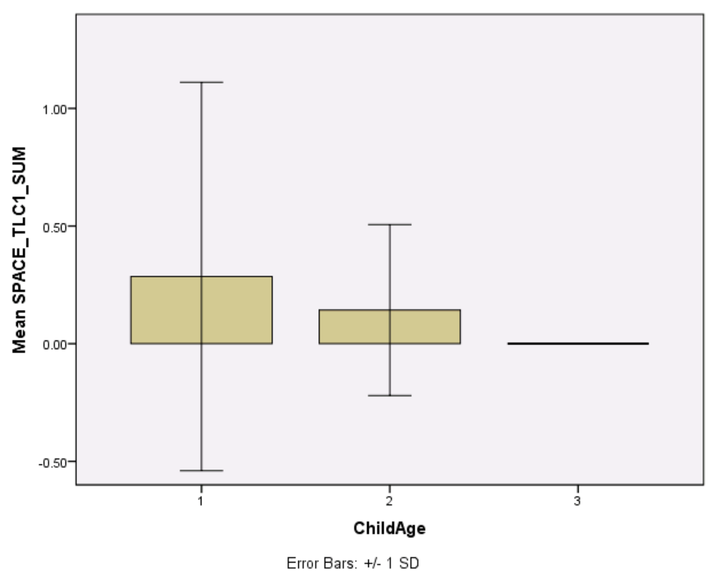


Figure 4.4.45. Total occurrences of losing tracking by aiming too close, per child, across age groups.

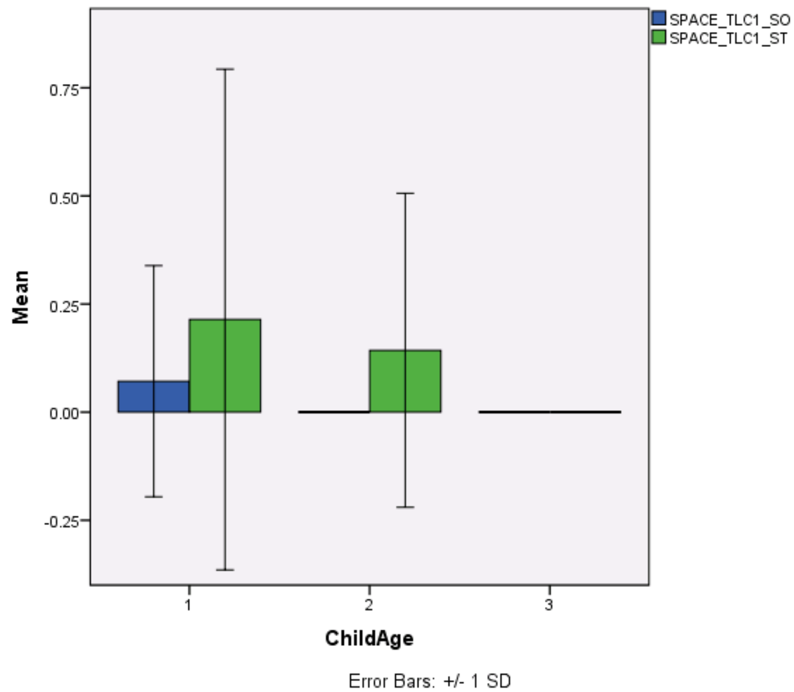


Figure 4.4.46. Total occurrences of losing tracking by aiming too close, per child, across No Tunnel levels (blue) and Tunnel levels (green)

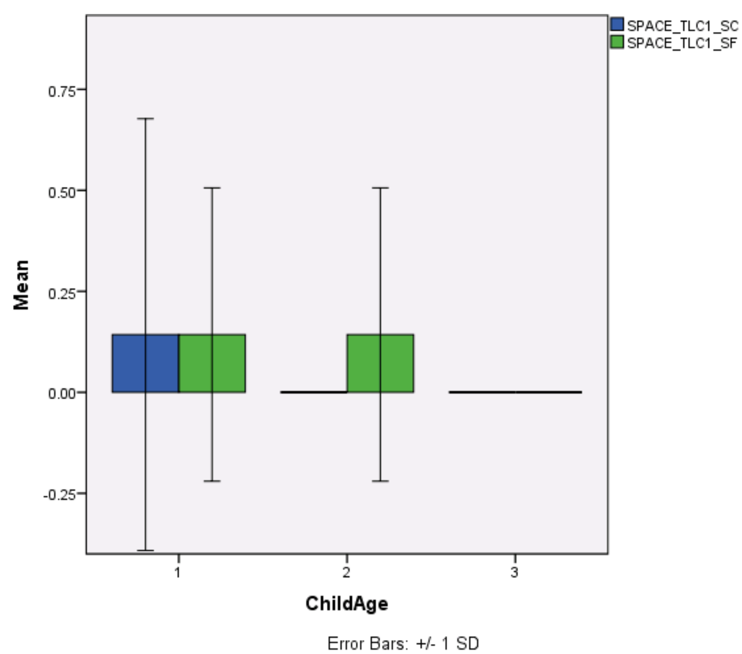


Figure 4.4.47. Total occurrences of losing tracking by aiming too close, per child, between Crosshair selection levels (blue) and Finger selection levels (green)

Difficulty orienting body in relation to the gameboard: Children sometimes have trouble orienting themselves around the gameboard. This issue was identified whenever children reported trouble, or required help, in relation to collecting lemons that were inside tunnels, or related to moving the phone in order to restore the augmented reality tracking. Below are examples of behaviors that were coded under this issue:

- (age 8) "I don't like this tube thing "
- (age 8) "It's hard to catch it" [talking about the level with tunnels and crosshair]
- (age 5) (experimenter gives verbal help) "It's easier if you take a step to the right to look inside the tunnel "
- (age 5) (experimenter gives physical help) [The experimenter suggests that the child takes a step back to recover tracking, but when the child doesn't move appropriately, has to take the phone in order to restore the tracking loss]

Table 4.4.12. Distribution of occurrences of difficulty orienting body in relation to gameboard.

	5-6 years old	7-8 years old	9-10 years old
Number of children experiencing this issue	9 (64%)	7 (50%)	0 (0%)
Number of occurrences per affected child Min – Max	1 - 4	1 - 2	NA
Number of occurrences per affected child Median	2	2	NA
Children experiencing this event at Severity 1 (frustration observed)	4 (29%)	7 (50%)	0 (0%)
Children experiencing this event at Severity 2 (verbal help given)	3 (21%)	1 (7%)	0 (0%)
Children experiencing this event at Severity 3 (experimenter had to move the child or phone)	6 (43%)	0 (0%)	0 (0%)

Only children in the 5-6 year old and 7-8 year old age groups were observed having these kinds of issues. When analyzing all children, **a statistically significant correlation was found after performing a two-way Pearson correlation between children's age (in months) and the number of observed number of orientation issues** ($r=-0.470$, $p=0.002$), indicating that as age increases, the number of such problematic events decreases. Also, an analysis of variance with between-subjects factor Age Group and within-subjects factors Selection Type and Movement Difficulty indicated **statistically significant effects: a main effect of Age Group** ($F(2,37)=6.13$, $p=0.005$), a **main effect of Movement Type** ($F(1,37)=20.16$, $p<0.001$), and an **interaction effect between Age Group and Movement Type** ($F(2,37)=6.13$, $p=0.005$). The interaction effect is explainable by the fact that children playing the No Tunnel levels did not experience any orientation issues, thus for that condition all age groups were the similar. For the Tunnels levels, post-hoc analysis indicated that the **Age Group factor shows a significant main effect** ($p=0.005$), and that significant differences exist between the 5-6 year olds and 9-10 year olds ($p=0.003$), but no significant differences exist between those groups and the 7-8 year olds.

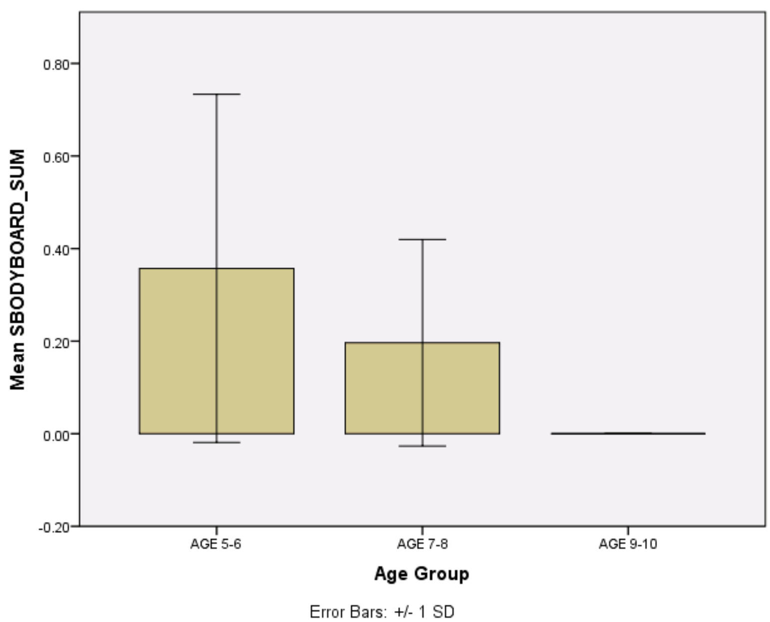


Figure 4.4.48. Total occurrences of difficulty orienting, per child, across age groups.

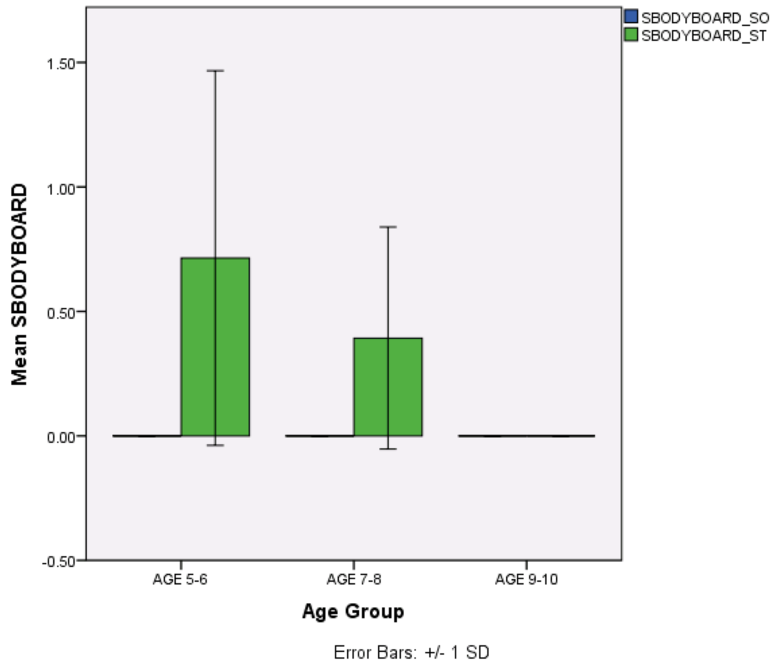


Figure 4.4.49. Total occurrences of difficulty orienting, per child, across No Tunnel levels (blue) and Tunnel levels (green)

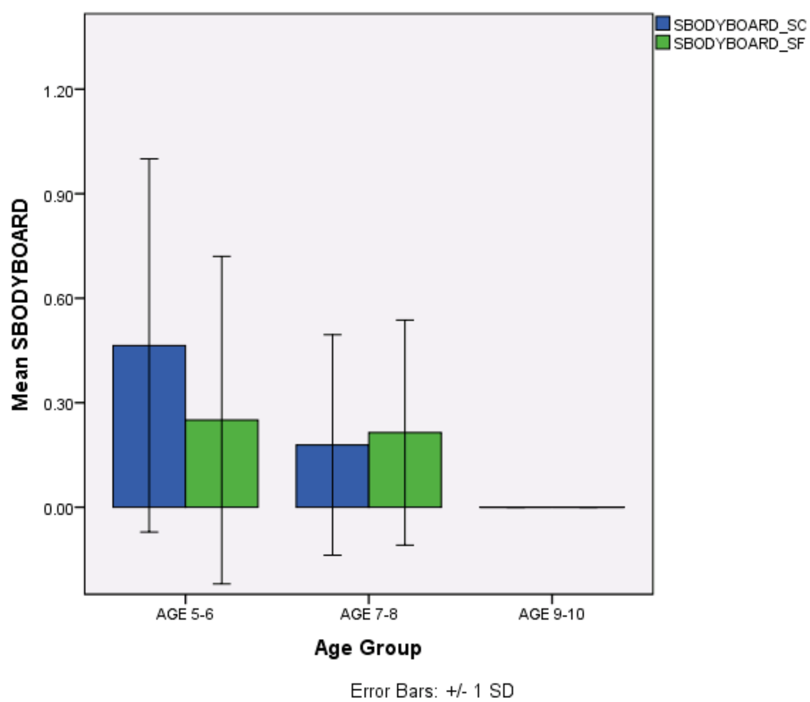


Figure 4.4.50. Total occurrences of difficulty orienting, per child, between Crosshair selection levels (blue) and Finger selection levels (green)

Non-significant descriptive results: When looking at Finger vs. Crosshair, the average number of body orienting problems in Crosshair conditions appears to be higher than that for Finger conditions for the 5-6 year-old children. One reason why this may be occurring is that in Crosshair levels children must use their body orientation in order to properly aim with the center of the screen, thus if the interaction is difficult this leads to more orientation issues. Another explanation is that Crosshair interaction is more difficult to perform because it is a method of indirect control (since children need to orient the screen first and then select items, rather than directly touch items with their finger), and this may cause increased cognitive load which leads children to be less able to control their body orientation as they attempt to target items using the crosshair.

4.4.6.2.5 Problems Related to Abstract Thinking

Needing initial instruction on how to use crosshair: All children knew how to use their finger to touch on the screen items of interest; however, many children did not know how to use the crosshair interaction which required them to aim with the center of the screen and touch on the side buttons. We coded this issue during the tutorial level. When coding this issue, we only coded one occurrence of this event – it was coded if, after 10 seconds of being exposed to the Crosshair selection type, whether the child needed help in order to proceed.

Table 4.4.13. Distribution of occurrences of needing initial crosshair instruction.

	5-6 years old	7-8 years old	9-10 years old
Number of children experiencing this issue (all at Severity 2)	10 (71%)	11 (79%)	7 (58%)

When analyzing all children, no statistically significant correlation was found after performing a two-way Pearson correlation between children's age (in months) and the occurrence of this event.

Non-significant descriptive results: Descriptive statistics indicate that, although this issue occurred in all age groups, less children in the 9-10 year old group required this kind of instruction.

Needing in-game instruction on how to use crosshair: During the tutorial level, all children were exposed to the Crosshair selection type and had to collect several lemons using this interaction. During the regular gameplay there were 2 levels requiring selecting lemons using the crosshair. In these levels, we observed some children which either had forgotten or did not realize that they needed to use the crosshair, and thus were provided verbal help.

Table 4.4.14. Distribution of occurrences of needing in-game crosshair instructions.

	5-6 years old	7-8 years old	9-10 years old
Number of children experiencing this issue (all at Severity 2)	2 (14%)	1 (7%)	0 (0%)
Number of occurrences per child Min - Max	1	1	NA

When analyzing all children, no statistically significant correlation was found after performing a two-way Pearson correlation between children's age (in months) and the occurrence of this event. Also, an analysis of variance indicated no statistically significant effect of between-subjects factor Age Group and within-subjects factors Selection Type and Movement Difficulty.

Non-significant descriptive results: Descriptive statistics indicate that this effect occurred for the two 5-6 year olds in the first Crosshair level (without tunnels), potentially indicating that they had forgotten how to use the interaction. For the 7-8 year old child this issue occurred in the second Crosshair level (with tunnels), which can indicate that the child knew how to use the interaction, since they had played the previous level correctly, but did not yet realize that the level required the same interaction.

Not understanding the game storyline: Children sometimes expressed confusion about why certain things happen in the fantasy world of the game. Knowing the answers to these questions may influence the children's motivation to engage with the game. For example:

- (age 7) "How can you make things out of lemons?"
- (age 9) "Why do you need to collect the lemons?"
- (age 8) "Why are the lemons in the tubes?"
- (age 8) "The imp.. makes green magic.. maybe that's why the tubes are green"

Table 4.4.15. Distribution of occurrences of being confused about the game storyline.

	5-6 years old	7-8 years old	9-10 years old
Number of children experiencing this issue (all at Severity 0)	1 (7%)	3 (21%)	1 (8%)
Number of occurrences per affected child Min – Max	1	1-2	1
Number of occurrences per affected child Median	1	1	1

When analyzing all children, no statistically significant correlation was found after performing a two-way Pearson correlation between children's age (in months) and the occurrence of this event. Also, an analysis of variance showed no statistically significant effects of between-subjects factor Age Group and within-subjects factors Selection Type and Movement Difficulty.

Non-significant descriptive results: Descriptive statistics indicate that this effect occurred more often in 7-8 year old children; it is unclear if the real population of 7-8 year olds is more concerned about game storyline in comparison to younger or older children. If so, one possible explanation may be that younger accept fantasy of the game more readily, or, they experience a higher degree of cognitive load while playing the game thus verbalize less concerns; and, older children may not question the game story because they are aware it is a fictional creation with limited logic.

Not understanding general game mechanics: Similarly, children expressed concerns about certain game mechanics, for example:

- (age 6) "So you just have to do four?" [the child is talking about collecting four lemons at a time]
- (age 7) "The time goes slow on this right?"
- (age 8) "So in the middle of each one there's like a temple or something? … that you have to collect the lemons on?"
- (age 6) "Why is it not giving me a time limit right now?... why is it not ticking right now?" [not seeing the timer tick]

Children had questions about how many lemons they had to collect, where they had to drop them off after collecting, and how/whether the timer functioned. Understanding the answers to these questions can lead to improved performance because the child knows what to do and expect from the game.

Table 4.4.16. Distribution of occurrences of being confused about general game mechanics.

	5-6 years old	7-8 years old	9-10 years old
Number of children experiencing this issue (all at Severity 0)	3 (21%)	2 (14%)	0 (0%)
Number of occurrences per affected child Min – Max	1-2	1-2	NA
Number of occurrences per affected child Median	1	1.5	NA

When analyzing all children, no statistically significant correlation was found after performing a two-way Pearson correlation between children's age (in months) and the occurrence of this event. Also, an analysis of variance no statistically significant effects of between-subjects factor Age Group and within-subjects factors Selection Type and Movement Difficulty.

Non-significant descriptive results: Descriptive statistics indicate that this effect occurred slightly more often in 5-6 year old children than in 7-8 year old children, and did not occur in 9-10 year old children. It may be that older children understood the game mechanics more, and/or they were less likely to verbalize their thoughts.

Difficulties interpreting tracking loss and recovering tracking: Losing the tracking of the paper is a significant issue because it interrupts the gameplay until the child recovers tracking. In the tutorial level, children were instructed and shown that the game works only when the phone looks at the gameboard through the front-facing camera positioned on the top-left corner of the phone. A degree of abstract thinking is required in order to understand why tracking has become lost and how to fix it. Under this usability issue, events were clustered which indicate that children either do not understand or are frustrated when the game loses tracking, and/or they needed help in order to make the tracking work again. Example events captured:

- (age 7) "Stop saying where is the paper! annoying"
- (age 7) Experimenter: "Your fingers are in the way"
- (age 9) [Trying to recover tracking] "OK.. it's facing the paper?"
- (age 6) "Paper's right there!!"
- (age 5) Experimenter: "It doesn't work if you are there you have to move a little bit" [the child is too far back]
- (age 5) Experimenter: "When it stops working you should come here on this side" [the child cannot fix the game so the experimenter has to move the phone]

When analyzing all children, a statistically significant correlation was found after performing a two-way Pearson correlation between children's age (in months) and the occurrence of this event ($r=-0.328$, $p=0.039$), indicating that as age increases, the occurrence of this event decreases. The average number of such events decreases with age. One reason is simply because the occurrence of tracking losses decreases with age in general. Another reason may be because older children play from a higher angle, and this leads the Vuforia tracking technology to recover tracking more easily. Finally, another possible explanation is that children become more aware of how to use the phone within the constraints of the technology, and they become more skilled at fixing the tracking loss.

Table 4.4.17. Distribution of occurrences of difficulties interpreting tracking loss and recovery.

	5-6 years old	7-8 years old	9-10 years old
Number of children experiencing this issue	7 (50%)	2 (14%)	2 (16%)
Number of occurrences per affected child Min – Max	1-6	1-5	1-1
Number of occurrences per affected child Median	2	3	1
Children experiencing this event at Severity 0 (confusion observed)	0 (0%)	0 (0%)	1 (7%)
Children experiencing this event at Severity 1 (frustration observed)	2 (14%)	1 (7%)	1 (7%)
Children experiencing this event at Severity 2 (verbal help given)	3 (21%)	2 (14%)	0 (0%)
Children experiencing this event at Severity 3 (experimenter had to move the child or phone)	5 (36%)	0 (0%)	0 (0%)

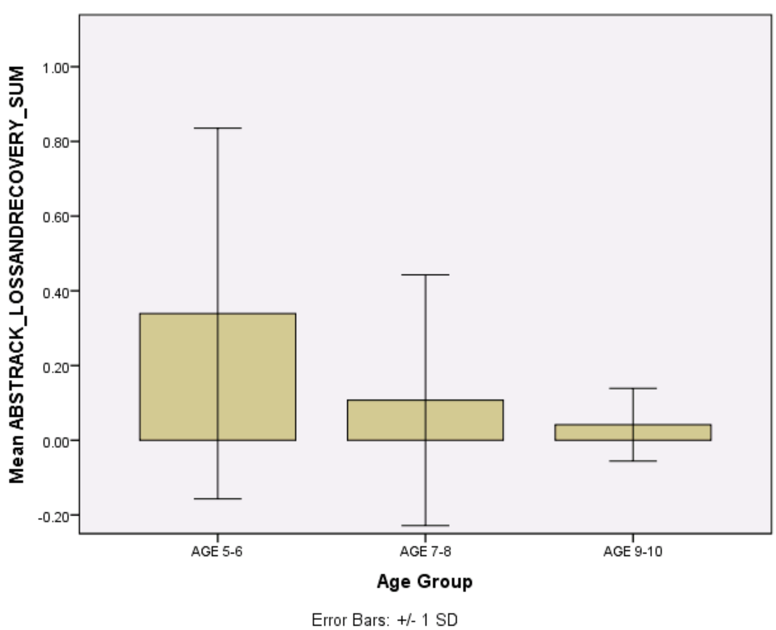


Figure 4.4.51. Total occurrences of difficulties interpreting tracking loss and recovery, per child, across age groups.

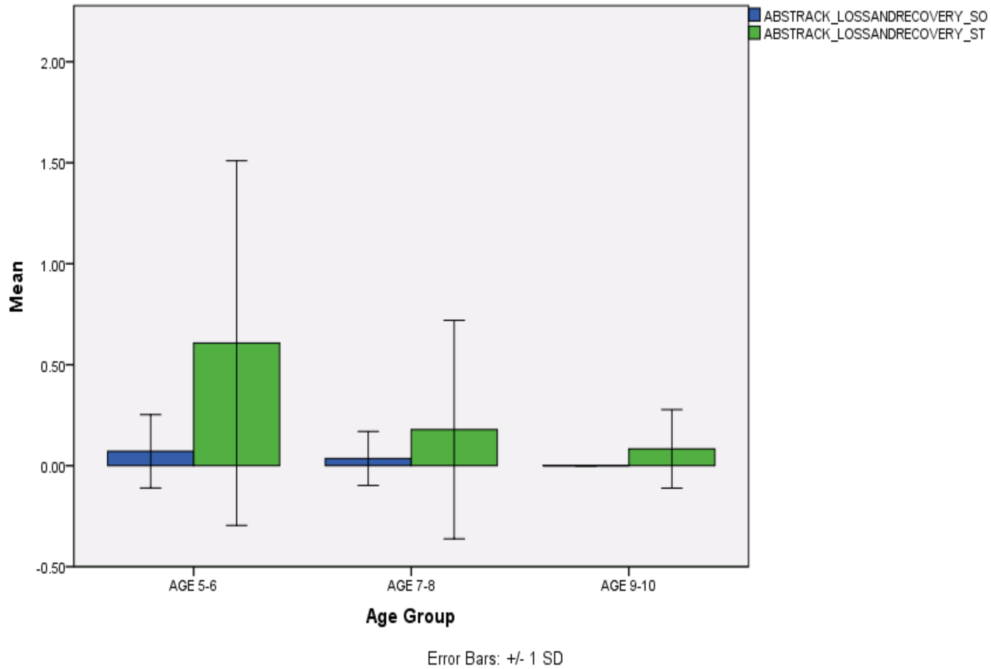


Figure 4.4.52. Total occurrences of difficulties interpreting tracking loss and recovery, per child, across No Tunnel levels (blue) and Tunnel levels (green)

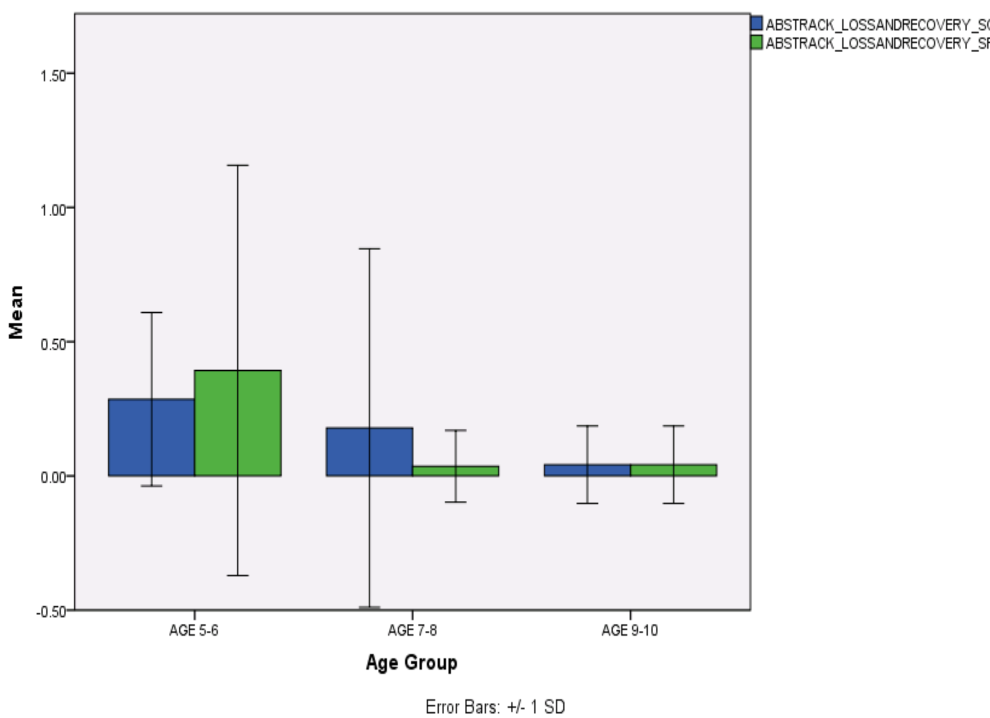


Figure 4.4.53. Total occurrences of difficulties interpreting tracking loss and recovery, per child, between Crosshair selection levels (blue) and Finger selection levels (green)

An analysis of variance with between-subjects factor Age Group and within-subjects factors Selection Type and Movement Difficulty indicated a **statistically significant main effect of Movement Difficulty** ($F(2,37)=8.00$, $p=0.007$), but **no statistically significant effect of Age Group or Selection Type.** This event occurred more often in Tunnel levels than compared to No Tunnel levels, because that is where most tracking losses occurred (see Section 4.4.3.2).

4.4.6.2.6 Problems Related to Attention

Bumping or tripping: Children of all ages were observed either bumping into the table or tripping over their own feet as they changed their perspective on the gameboard. This event may indicate a focused attention on the game, and may cause problems if children play such games in dynamic environments such as in busy classrooms or outdoors.

Table 4.4.18. Distribution of occurrences of bumping or tripping.

	5-6 years old	7-8 years old	9-10 years old
Number of children experiencing this issue (all at Severity 1)	5 (36%)	5 (36%)	5 (41%)
Number of occurrences per affected child Min – Max	1-2	1-3	2-3
Number of occurrences per affected child Median	1	1	2

When analyzing all children, no statistically significant correlation to child Age was found after performing a two-way Pearson correlation between children's age (in months) and the occurrence of this event. An analysis of variance with between-subjects factor Age Group and within-subjects factors Selection Type and Movement Difficulty indicated a statistically significant main effect of Movement Difficulty ($F(1,37)=15.74$, $p<0.001$); no other statistically significant effects were found. This effect occurred significantly more in Tunnel levels, which is expected since that is where children had to move their body in order to change perspective around the gameboard.

Non-significant descriptive results: The descriptive statistics indicate that, although this effect occurred across all age groups, older children experienced slightly more such events. Also, Finger conditions show slightly higher number of bumping than Crosshair conditions; this might be because Crosshair requires the child to slow down and focus on the selection task.

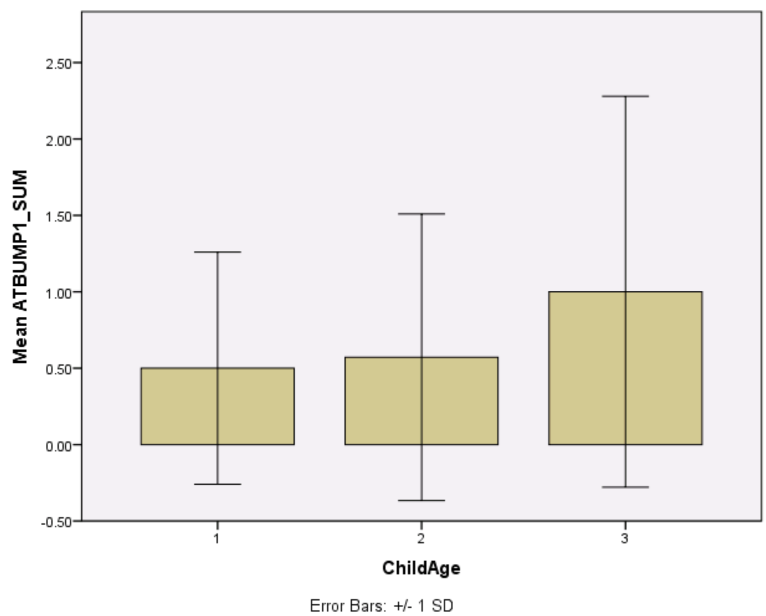


Figure 4.4.54. Total occurrences of bumping or tripping, per child, across age groups.

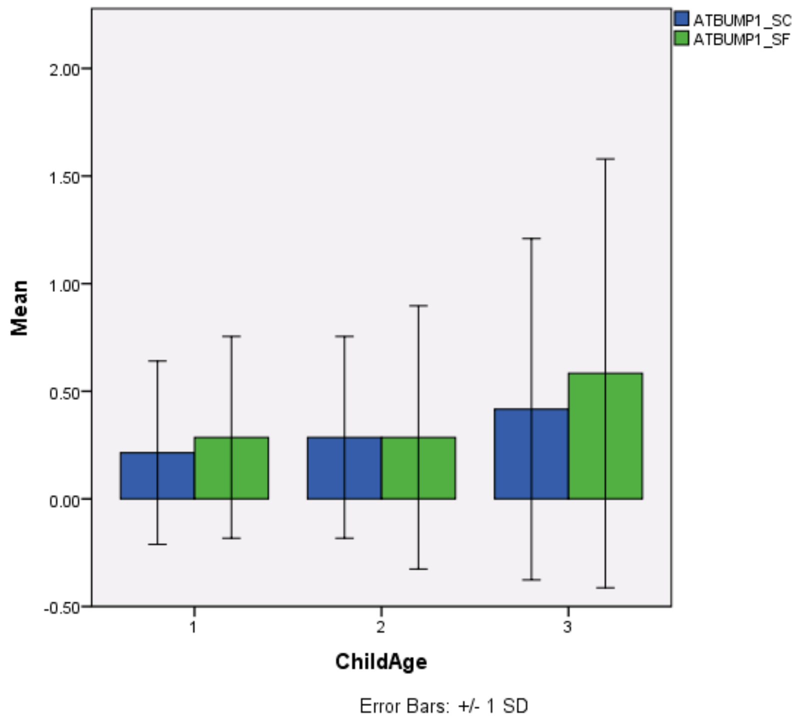


Figure 4.4.55. Total occurrences of bumping or tripping, per child, between Crosshair selection levels (blue) and Finger selection levels (green)

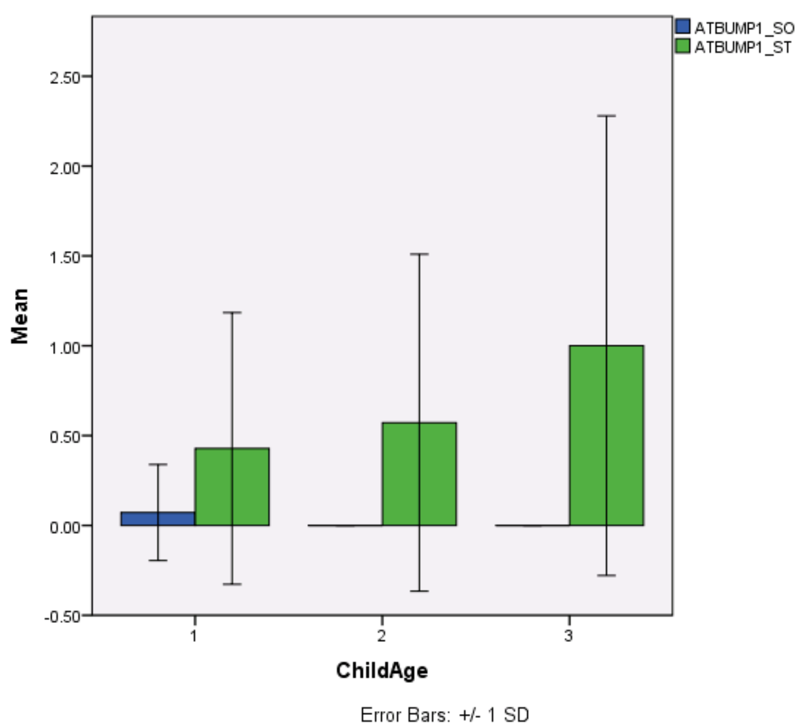


Figure 4.4.56. Total occurrences of bumping or tripping, per child, across No Tunnel levels (blue) and Tunnel levels (green)

Interruption due to self-distraction: Children were sometimes observed losing interest in the game and either using the camera to look away at other pieces of the environment, or looking at other people such as the experimenter or the parent (who were positioned behind the child at all times during gameplay). This is not necessarily a usability problem, but it can indicate split attention, or boredom with the game.

Table 4.4.19. Distribution of occurrences of self-distracted interruptions.

	5-6 years old	7-8 years old	9-10 years old
Number of children experiencing this issue (all at Severity 0)	3 (21%)	1 (7%)	1 (8%)
Number of occurrences per affected child Min – Max	1-30	8-8	6-6
Number of occurrences per affected child Median	2	8	6

When analyzing all children, no statistically significant correlation was found after performing a two-way Pearson correlation between children's age (in months) and the occurrence of this event. Also, an analysis of variance no statistically significant effects of between-subjects factor Age Group and within-subjects factors Selection Type and Movement Difficulty.

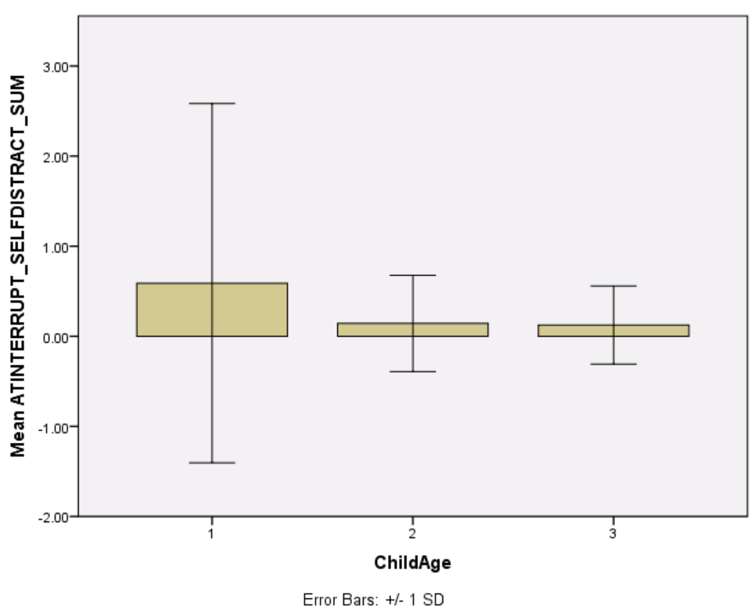


Figure 4.4.57. Total occurrences of self-distracted interruptions, per child, across age groups.

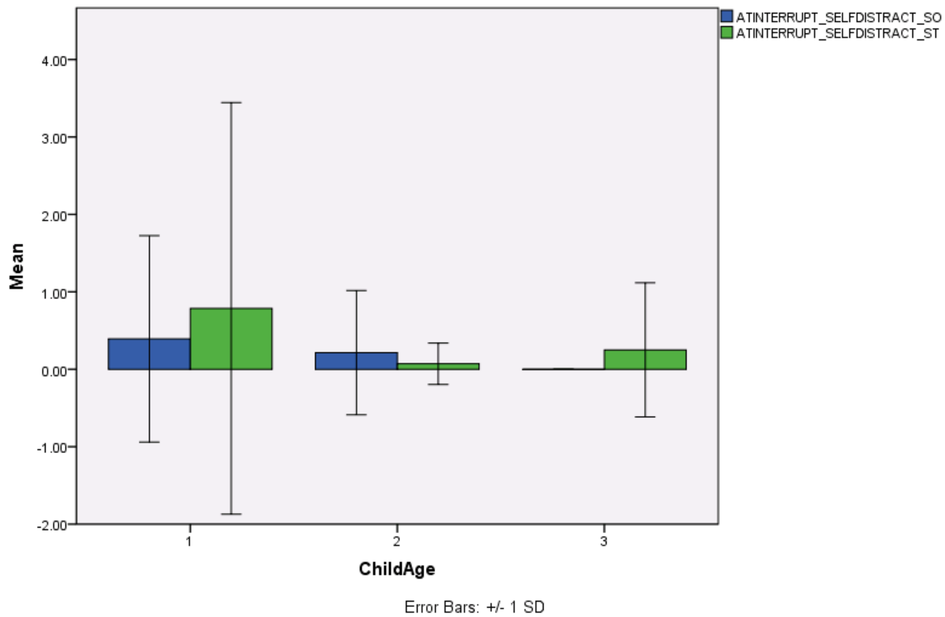


Figure 4.4.58. Total occurrences of self-distracted interruptions, per child, across No Tunnel levels (blue) and Tunnel levels (green)

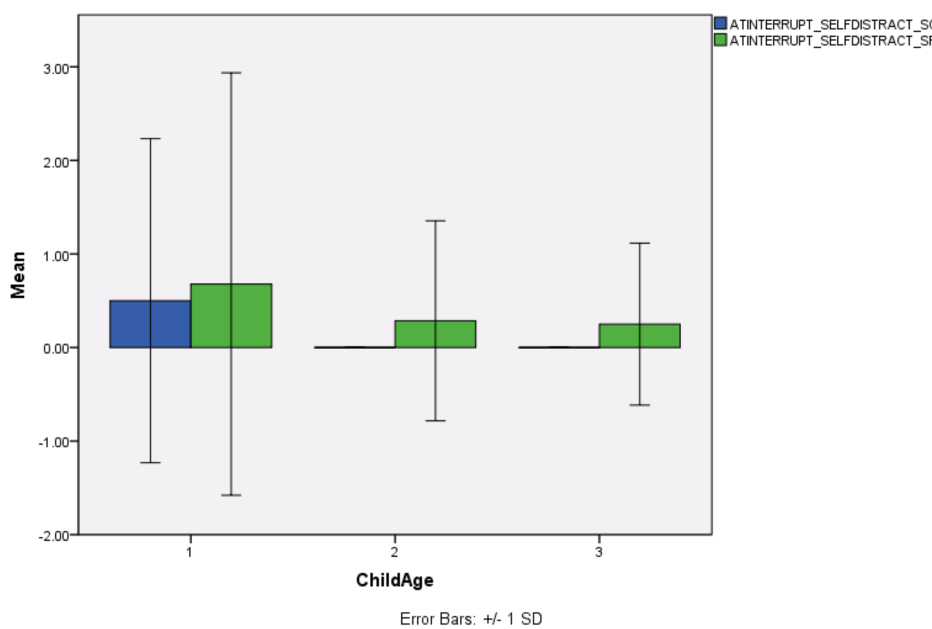


Figure 4.4.59. Total occurrences of self-distracted interruptions, per child, between Crosshair selection levels (blue) and Finger selection levels (green)

Non-significant descriptive results: Descriptive statistics indicate that this behavior occurred in more of the 5-6 year olds than the other age groups. In younger children this behavior occurred in both Finger and Crosshair interactions, and in older children the effect occurred in Finger interactions. This might indicate that Finger interactions are less challenging in terms of attention; however, no conclusions can be drawn based on the low amount of occurrences for this behavior.

Interruption due to scratching: Children were observed to interrupt their gameplay in order to scratch themselves. Like the previous interruption due to self distraction, this is not necessarily a usability problem, but it can indicate split attention or boredom with the game.

Table 4.4.20. Distribution of occurrences of scratching interruptions.

	5-6 years old	7-8 years old	9-10 years old
Number of children experiencing this issue (all at Severity 0)	3 (21%)	8 (57%)	8 (67%)
Number of occurrences per affected child Min – Max	1-8	1-3	1-12
Number of occurrences per affected child Median	1	2	3

When analyzing all children, a statistically significant Spearman correlation found a positive association between children's age (in months) and the occurrence of this event (Rs=0.332, p=0.036). It is possible that scratching is an indication of low cognitive load possibly related to older children being more comfortable holding the phone with one hand and finding the game easier to play. It is also possible that scratching is an indication of low motivation, which may be possible because the game was designed to motivate younger children; however, this may not be the case, since self-reported fun was not statistically different between age groups. No statistically significant effects were found for Selection Type or Movement Difficulty.

A Spearman correlation test detected **a significant positive correlation between interruption due to scratching and the use of No Handed grip in either hand**, after the effect of Age was removed from both factors via linear regression (Rs=0.467, p=0.002). It appears that scratching occurs more often children who are comfortable holding the phone with one hand.

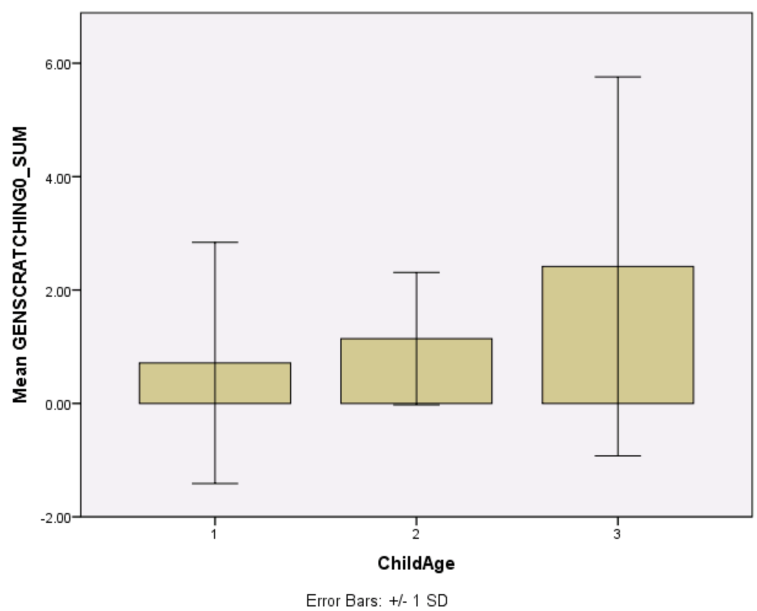


Figure 4.4.60. Total occurrences of scratching interruptions, per child, across age groups.

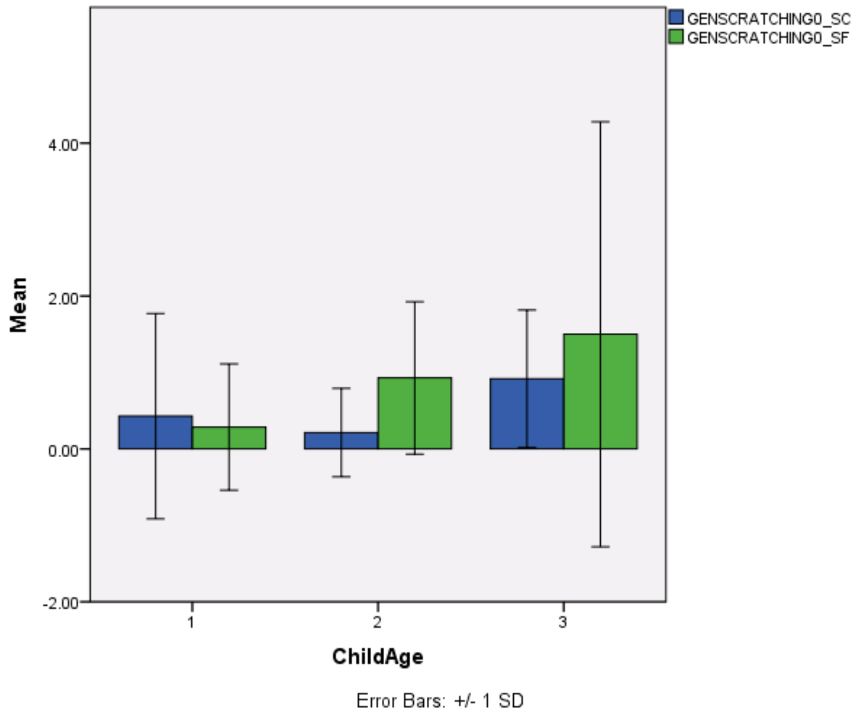


Figure 4.4.61. Total occurrences of scratching interruptions, per child, between Crosshair selection levels (blue) and Finger selection levels (green)

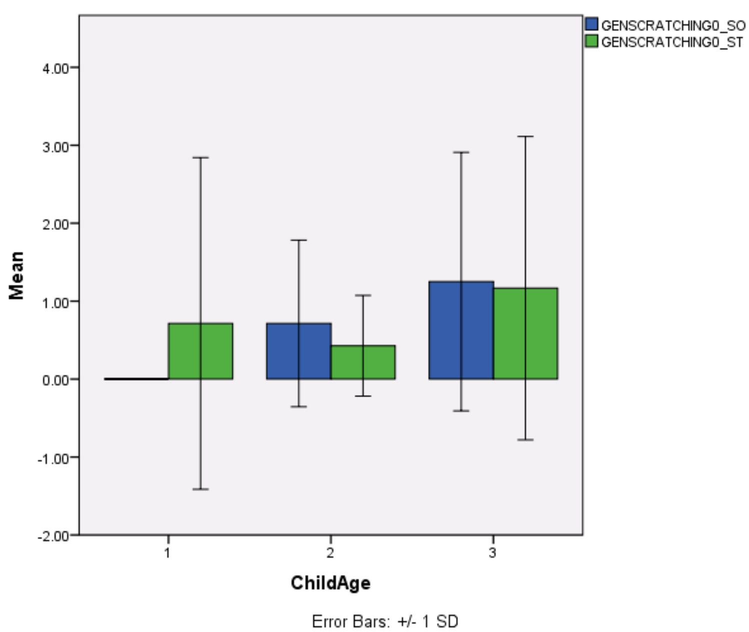


Figure 4.4.62. Total occurrences of scratching interruptions, per child, across No Tunnel levels (blue) and Tunnel levels (green)

Non-significant descriptive results: In 5-6 year olds this behavior occurred only in Tunnel levels, while in older children it appears to occur slightly less in Tunnel levels. Scratching tends to occur more often in Finger than Crosshair levels. One reason may be that Crosshair requires more focus. Or, it may be that children use a different grip for Finger interactions, which frees up their hands for scratching.

4.5 EMERGING FACTOR ANALYSES

During the study, several factors emerged as potentially influencing children's performance and usability problems. These factors have been investigated even though they were not included in the original research questions. The sections below present the analysis of effects of grip, practice, technology experience, and gender.

4.5.1 Analysis of Grips Used by Children

4.5.1.1 Types of Hand Grips

When children held the phone they used a variety of grips. During video coding, the behavior of each hand was tagged according to the grip being used. Grips were defined according to the direction of the forces for holding/manipulating the phone. The grips were defined as follows:

CRAB GRIP: The phone is clamped by the hand as if in a vice, and the force holding the phone is applied from the top of the phone and from the bottom of the phone. When children hold this grip they usually use the thumb to hold the bottom of the phone, and they use one or more of their fingers to hold the top of the phone.

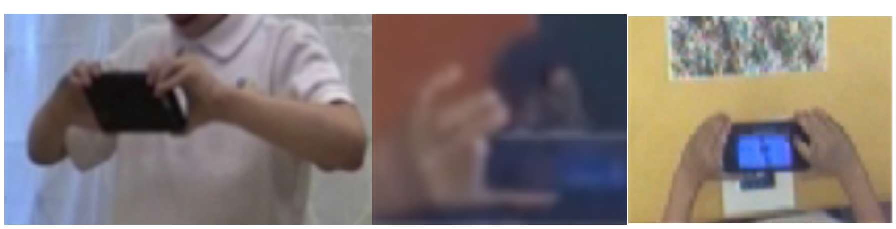


Figure 4.5.1. Examples of Crab grips.

STRAIGHT GRIP: The phone is held by applying force to only on the side of the phone. This typically occurs when children hold their palm flat on the side of the phone. In order to hold this grip, the hand must be pushing on the phone from the side, and an opposite force must be applied by the other hand to stabilize the phone. If the hand curls and touches the phone, then force is applied from the back of the phone and therefore the grip becomes a Curl grip.

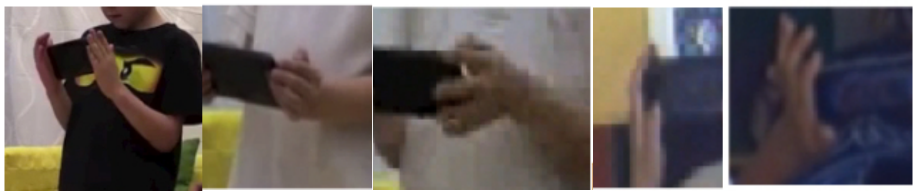


Figure 4.5.2. Examples of Straight grips.

CURL GRIP: The phone is held by the hand in a vice, where the force is applied from the side and also from the back of the phone. This grip generally looks as if the hands are curled around the phone.

Figure 4.5.3. Examples of Curl grips.

CORNER GRIP: The phone is held from a corner on the bottom of the phone. If the phone is held by 3-or-less fingers positioned in a corner, then it's considered a corner grip because the forces are localized to the corner of the phone. If there are 4 or more fingers, then it is a Curl grip because there is a significant force from the side and back of the phone.

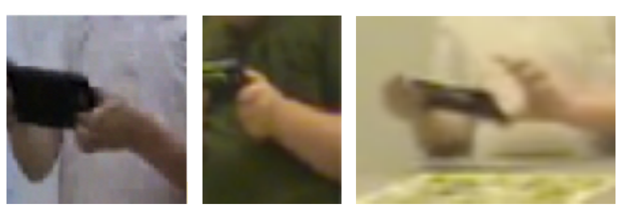


Figure 4.5.4. Examples of Corner grips.

BOTTOM GRIP: The phone is resting on at least one of the hands, located on the bottom edge of the phone but not in a corner. Typically, the palm of the hand is used to apply upward force on the bottom of the phone, while the fingers are supporting the back of the phone.

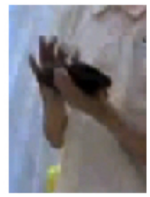


Figure 4.5.5. Example of Bottom grip (on the child's left hand).

NO GRIP: The hand is not used to support the phone.

Figure 4.5.6. Examples of No Grip (on the child's right hand)

4.5.1.2 Analysis of Grips

The descriptive statistics for children's use of grips are found in Appendix C. In order to understand children's use of different grips, I first analyzed the grips in each hand separately from each other. The phone camera was placed on the top left corner of the Atrix HD smartphone, thus strongly influencing the child's grip in the left hand, and potentially causing different styles of grips in each hand. In order to determine if there are statistically significant differences between use of grips within each hand, I first calculated the percentage of time that a child used each grip within each level, then performed an ANOVA analysis with between-subjects factor Age Group, and within-subjects factors Selection Type, Movement Type, and Grip Type.

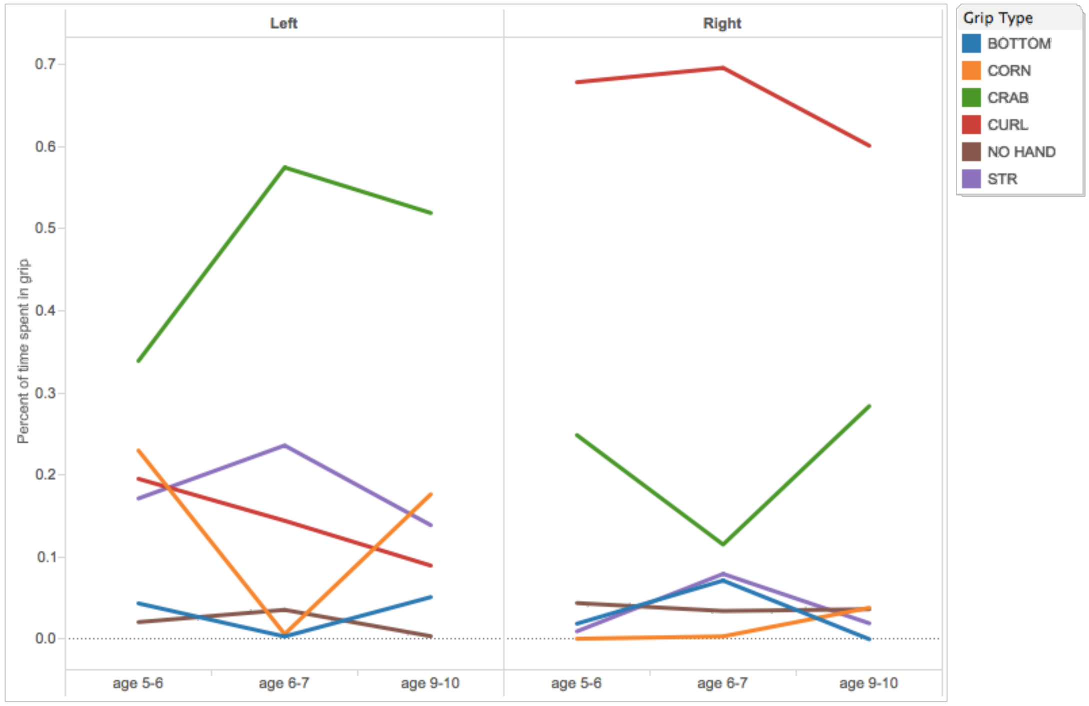


Figure 4.5.7. Percentage of time spent in each grip, in the Left and Right hands.

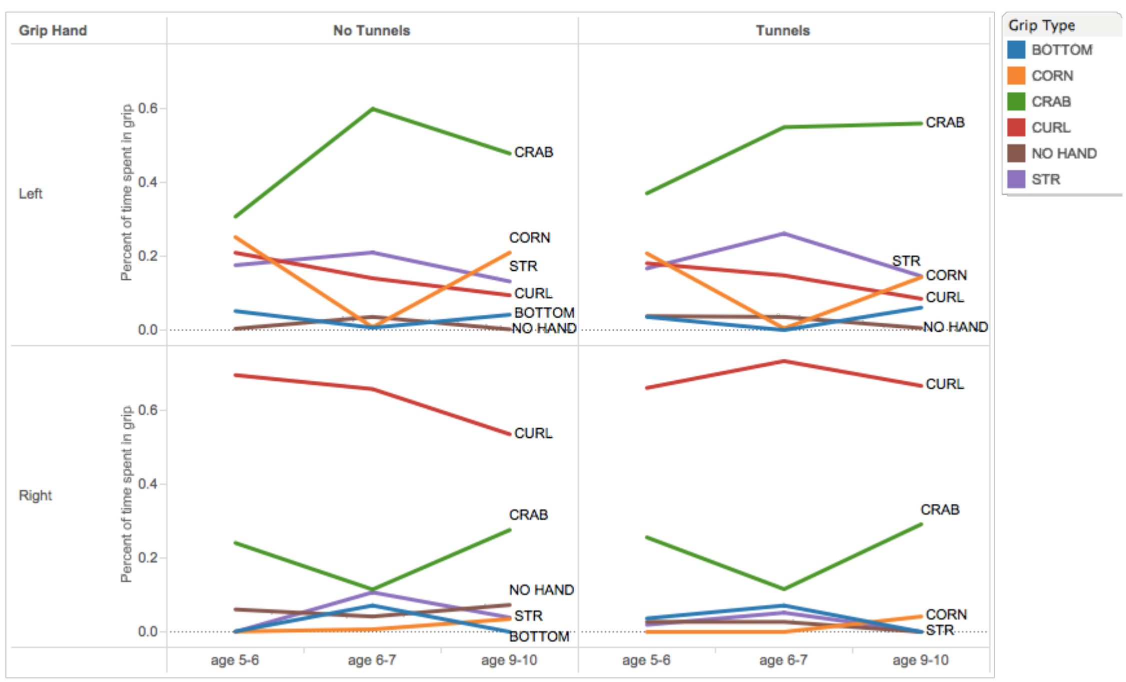


Figure 4.5.8. Percentage of time spent in each grip, in the left (top) and right hand (bottom), between No Tunnels (left) and Tunnels (right) conditions.

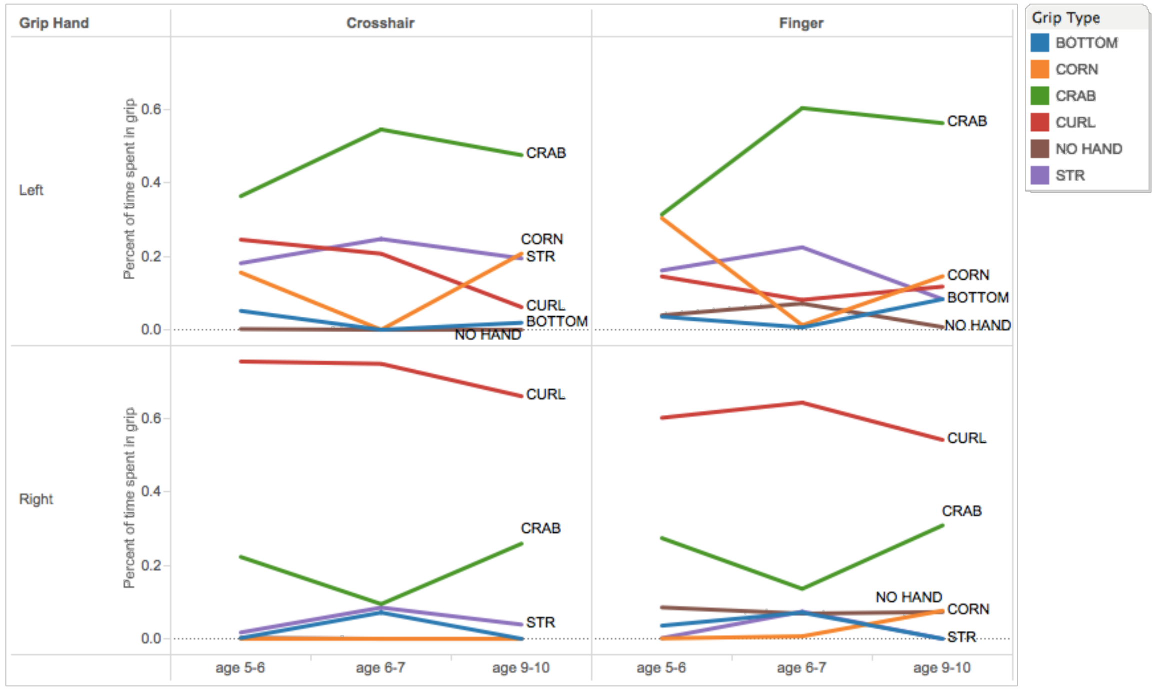


Figure 4.5.9 . Percentage of time spent in each grip, in the left (top) and right hand (bottom), between Crosshair (left) and Finger (right) conditions.

For the left hand, there was a significant main effect of Grip Type ($F(5,33)=21.67$, $p<0.001$); no statistical main effects or interaction effects were found involving Age Group, Selection Type, or Movement Difficulty. Based on the descriptive statistics it is visible that the CRAB grip is most popular overall, followed by CURL, STRAIGHT and BOTTOM. Post-hoc Bonferroni-corrected contrasts were conducted to investigate the differences between the grips of CRAB, CURL, STRAIGHT and BOTTOM. Only 5 statistical tests were conducted, in order to use a per-test Type I error threshold of 0.01. The tests indicate that the CRAB grip is used significantly longer than the average of all other grips ($p<0.001$); furthermore, the CRAB grip is used significantly longer than compared to the CURL grip ($p=0.001$), and significantly longer compared to the STRAIGHT grip ($p=0.006$). However, the CURL and STARIGHT grips are not significantly different from each other ($p=0.561$). The BOTTOM grip is not significantly different than the average of the CURL and STRAIGHT grips ($p=0.605$).

Non-significant descriptive results: Based on the descriptive statistics of left hand grips (listed in Appendix C), some trends are visible across age groups. Young children have a low usage of each grip type (on average, no grip is used more than 35% of the time). In middle-age children, there is a strong preference for CRAB grips and an increased preference for STRAIGHT grip, while the use of CURL grip and CORNER grips decreases. Finally, in older children, there still appears a strong preference for CRAB, with an increased preference for CORNER and BOTTOM grips, while the use of STRAIGHT and CURL grips decreases.

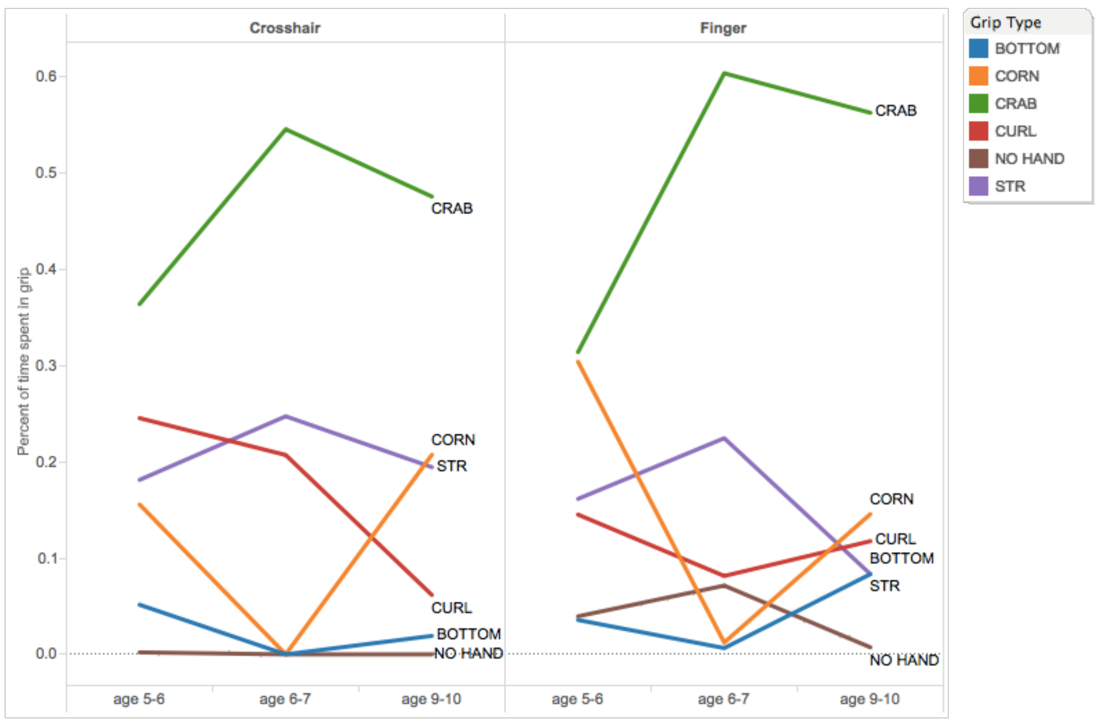


Figure 4.5.10. Percentage of time spent in each grip in the left hand, in Crosshair (left) and Finger (right) conditions.

On the other hand, for the right hand there was **a significant main effect of Grip Type** ($F(5,33)=46.49$, $p<0.001$), as well as **a significant interaction effect of Grip Type with Selection Type** ($F(5,33)=2.79$, $p=0.033$). The interaction effect between Grip Type and Selection Type was analyzed first, investigating significant differences between Finger and Crosshair conditions in relation to the grips that appeared to differ most between the two conditions: CURL, CRAB, STRAIGHT, CORNER and NO-HAND . Post-hoc Bonferroni-corrected contrasts were conducted, and only 5 statistical tests were conducted in order to preserve a low Type I error rate of 0.01. There were **no statistical significant differences observed between each grip type within the two Finger vs. Crosshair conditions**. The tests indicate that the percentage of grip use between the Finger and Crosshair conditions is not statistically different for the CURL grip (p=0.020), NO-GRIP (p=0.023), CORNER grip (p=0.271), CRAB grip (p=0.377), and STRAIGHT grip (p=0.412). Based on the descriptive statistics comparing between Finger and Crosshair levels on right hand, it appears that the CURL grip is used more often in Crosshair levels by all age groups; to a much lesser degree, the Crosshair levels show a lower use of NO GRIP and CORNER grips.

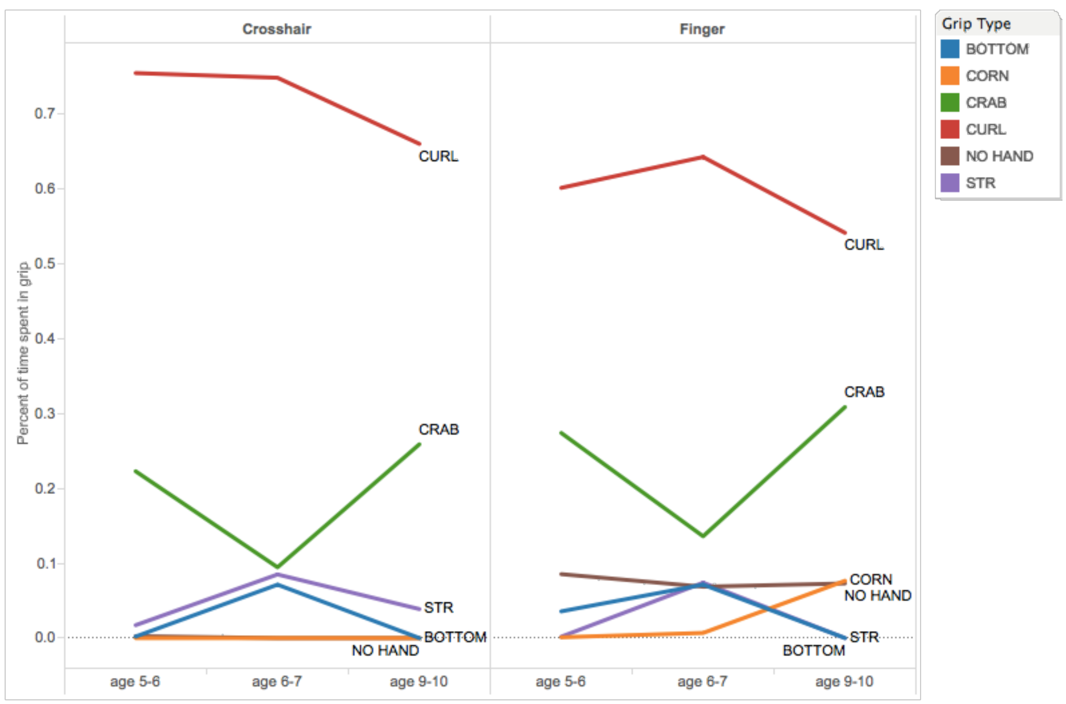


Figure 4.5.11. Percentage of time spent in each grip in the right hand, in Crosshair (left) and Finger (right) conditions.

When analyzing the significant main effect of Grip Type aggregated over all gameplay conditions, I performed post-hoc Bonferroni-corrected contrasts to investigate the differences between the most popular right-hand grips (CURL, CRAB, STRAIGHT, NO-GRIP, and BOTTOM). Only 5 statistical tests were conducted, in order to use a per-test Type I error threshold of 0.01. The tests indicate that **the CURL grip is used significantly longer than CRAB grip** ($p<0.001$). The **use of the CRAB grip is also significantly larger than compared to the average use of all other grips** ($p=0.002$), and **compared to the use of the NO-GRIP grip** ($p<0.001$). There were **no statistically significant differences in grip use detected when comparing NO-GRIP to STRAIGHT grip** ($p=0.96$), or **when comparing STRAIGHT grip to BOTTOM grip** ($p=0.88$), indicating that the use of these other grips may not be different.

Non-significant descriptive results: Based on the descriptive statistics of right hand grips (listed in Appendix C), some trends are visible across age groups. The use of the CURL grip is consistently high across all age groups, however showing an increase in middle children and a decrease for older children; the CRAB grip is also popular across all age groups, although much less than CURL, with middle children using it less, and older children using it most. Compared to other age groups, middle-aged children show increase preference for CURL, decreased preference for CRAB, and slightly increased preference for BOTTOM and STRAIGHT grips.

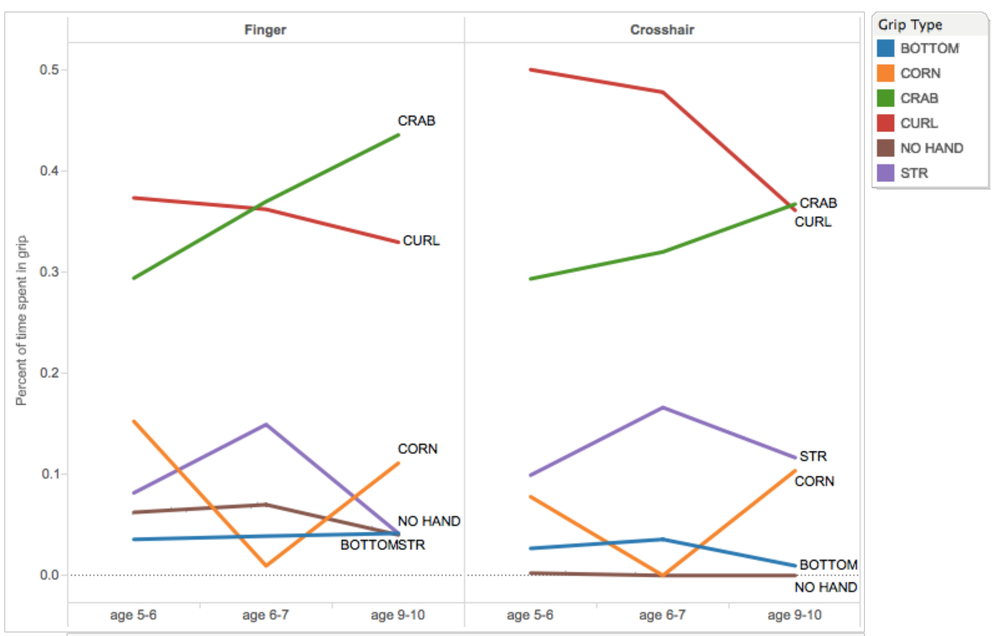


Figure 4.5.12 Percentage of time used for each grip in both hands while using the Finger selection (left) or Crosshair selection (right).

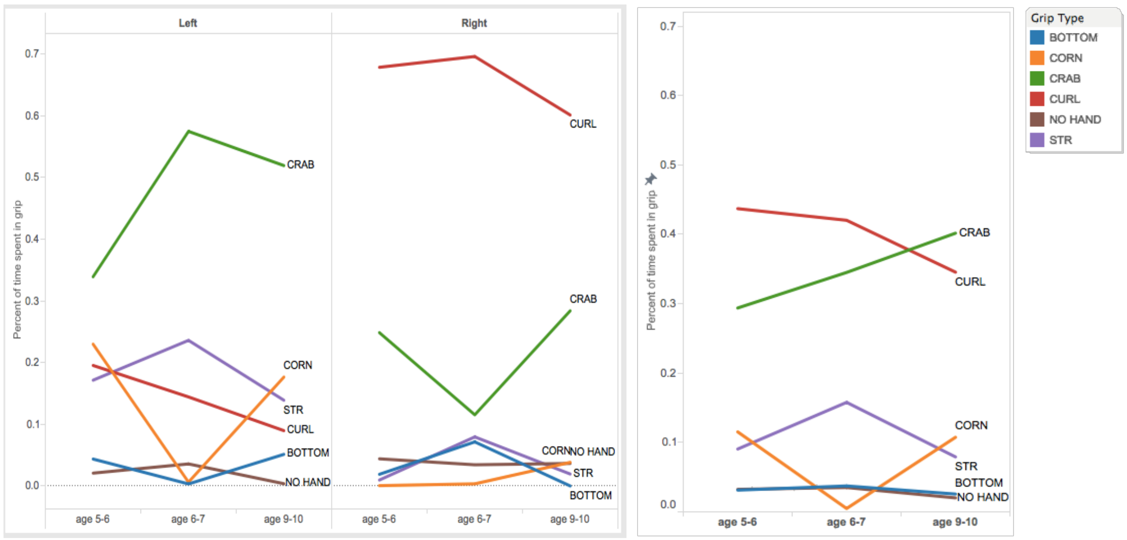


Figure 4.5.13. Percentage of time used for each grip in the left hand (left), right hand (middle), and overall (right).

From the descriptive statistics between both hands, there are interesting trends to be observed. Children 5-6 years old use CURL and CRAB grips most often. The CRAB grip is strongly used in the left hand, and shows increasing popularity overall between age groups; it becomes a popular grip in the left hand starting in the 7-8 year old group (although this grip shows a decrease in CRAB use for the right hand) and continues to increase in the 9-10 year old group. The CURL grip shows decreased usage as children become older, although it is used strongly in the right hand by all age groups. The STRAIGHT and CORNER grips show a low amount of use, mostly in the left hand, and their use appears similar between 5-6 and 9-10 year old children; for 7-8 year olds, they seem to prefer STRAIGHT grip over CORNER for the left hand. The BOTTOM and NO GRIP are used least often by all age groups, with a slight increase in the 7-8 year olds.

I also analyzed the relationships between grips and AR performance and usability problems encountered by children. As reported in the previous sections, only two significant correlations were found after accounting for the effects of age. The occurrence of children **losing**

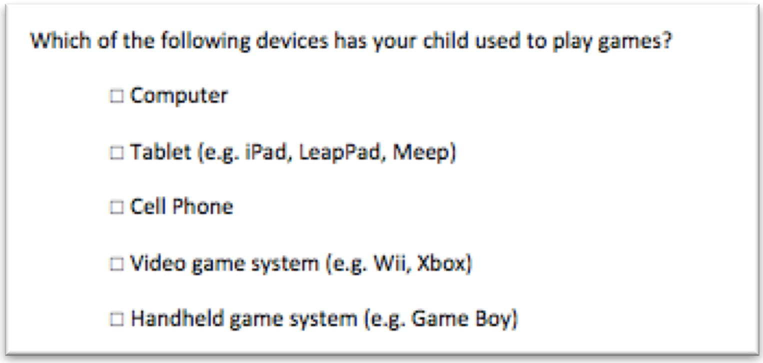
tracking due to finger in the way was significantly correlated positively to the use of "Straight" grip in the left hand. The phone camera is on the left side of the phone, thus this grip is likely causing the tracking loss. Furthermore, the occurrence of **children interrupting the gameplay to scratch themselves was significantly positively correlated to the use of "No Handed" grip in either hand**. The causal relationship between these events is unclear. When a child needs to scratch themselves, they may choose to hold the phone with one hand, thus these two events are likely to occur together; however, it is unclear if a willingness to scratch will occur more often in children who are more skilled at collecting items while holding the phone with one hand.

4.5.2 Analysis of Speed vs. Accuracy Effects

I analyzed the correlation between completion time and selection accuracy, in order to determine if children were attentive to the task, or whether they were rushing. I expected to see a negative correlation if children were not paying attention (whereby they would collect lemons quickly but have more errors), and a positive correlation if they were paying attention (if they would be attentive to each selection event, more time would pass as they made more errors). **A significant positive correlation was found overall between task time and selection errors** (R_s=0.431, p=0.004)**.** However, analyzing within age groups, a significant correlation was only detected **within the group of 7-8 year olds** (R_s=0.701, p=0.005).

4.5.3 Analysis of Previous Experience with Technology

In the questionnaire completed by the parent, I asked parents to indicate which devices children used to play games (Figure 4.5.14). This question was used as an indicator of children's exposure to technology.


Which of the following devices has your child used to play games?

☐ Computer

☐ Tablet (e.g. iPad, LeapPad, Meep)

☐ Cell Phone

☐ Video game system (e.g. Wii, Xbox)

☐ Handheld game system (e.g. Game Boy)

Figure 4.5.14. Parent survey question about child's exposure to technology.

In order to analyze if children's exposure to technology is correlated to AR performance, I performed two-way nonparametric Spearman correlations (since parametric assumption of normality was violated) between the number of technology items indicated on this question, and each of the AR performance metrics, after removing the effects of Age from both factors using linear regression. Unexpected **significant positive correlations were found between the number of technology devices and the number of selection errors** (Rs(42)=0.368, p=0.016), indicating that more exposure to technology is associated with higher inaccuracy. Further unexpected results are that **significant positive correlations were found between the number of technology devices and the number of tracking losses** (Rs(42)=0.511, p=0.001), indicating that more exposure to technology is associated with higher likelyhood of children losing AR tracking.

I then investigated whether there is a difference between males and females in terms of technology experience. Nonparametric Mann-Whitney U test showed significant differences, whereby **technology experience for boys was significantly higher than for girls** (mean rank girls = 17.28, boys = 25.34, U=304.50, p=0.026). The significant gender difference remained after accounting for the effects of age on both variables.

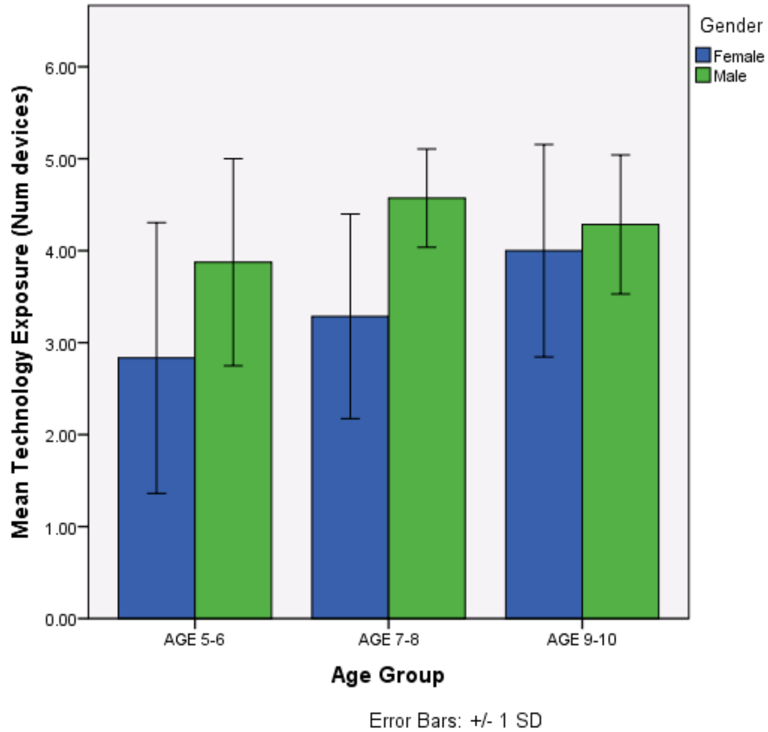


Figure 4.5.15. Average number of devices used, for Females vs Males.

The association between gender and technology experience indicates that gender may be a confounding factor in the above analysis, of technology experience vs. AR performance metrics. In order to investigate this, I performed the same two-way nonparametric Spearman correlations between children's previous technology experience and each of the AR performance metrics, but this time removing the effects of Age as well as Gender from both factors using linear regression**. Technology experience remained significantly correlated to number of tracking losses** (Rs(42)=0.462, p=0.002), but became marginally nonsignificant in relation to number of selection errors (Rs(42)=0.284, p=0.069). This change in results regarding number of selection errors indicates that the relationship between gender and technology experience is confounding the analysis; however, within the current experimental design I am unable to determine whether technology experience is causing the effect on errors, or if gender is causing the effect. This will be discussed as a limitation of the current study.

4.5.4 *Analysis of Gender Differences*

In order to analyze potential effects of gender, I first performed a mixed ANOVA with between-subjects factors Age and Gender, and within-subjects factors Movement Difficulty and Selection Type, for all the dependent measures of performance and usability problems. The data violated the normality assumption of the parametric test. Nonparametric analysis using Mann-Whitney U tests yielded the same significant results as the parametric analysis, thus the parametric analysis is reported here. **A significant main effect of Gender was found for number of selection errors**, whereby the average number of errors per lemon for girls (M=0.58, SD=0.39) had 48% less errors than boys.

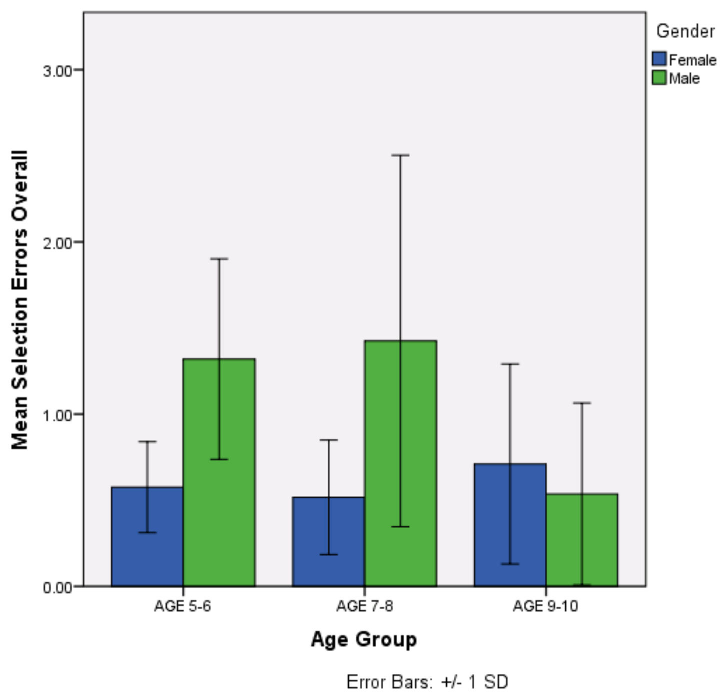


Figure 4.5.16. Average number of selection errors (per lemon), for Females vs Males.

However, the analysis of previous experience with technology (Section **Error! eference source not found.**) has shown that girls in this subject pool had significantly less technology experience than boys, and that technology was significantly correlated to some metrics of AR performance. The correlation between gender and technology experience may indicate that previous experience with technology is a confounding variable in the gender effects observed above. To investigate this, I performed nonparametric Mann-Whitney U tests on the effects of gender on AR performance metrics, after using linear regression to remove the effects of technology experience from the performance metrics. **No significant gender effects were found after accounting for technology experience**.

4.5.5 Analysis of Hand Size

In order to analyze potential effects of hand size, a two-way Spearman correlation was performed between Hand Size and each of the AR performance metrics under all experimental conditions, after removing the effects of Age (in months) from each factor. **No significant correlations were found ($p>0.05$).**

4.5.6 Analysis of Order and Practice Effects

Children in the study were exposed to 4 different levels of the game. The order of selection type was randomized: children were either assigned to experience Finger selection first followed by Crosshair selection, or vice-versa. The order of movement type was not randomized: within each of those selection type conditions, children always experienced a level with No Tunnels followed by a level with Tunnels.

I first analyzed whether the order of selection type exposure caused any differences in relation to each child's overall AR performance and usability problem metrics, tested through Mann-Whitney U tests. **No significant effects were found**, indicating that the random assignment successfully counterbalanced order effects.

I then investigated whether children's performance changed over time as they progressed through levels of the game. I was interested in whether performance on the first two levels played is different from the last two levels played – if the order effects are counterbalanced, and if children improve over time, their overall performance scores should improve between the first and second half of the game. I was also interested in whether the magnitude of differences between the No Tunnel levels compared to the magnitude of differences in the Tunnel levels; that is, if children improve more on the No Tunnel levels than compared to the Tunnel levels, or vice versa. For each AR performance metric and usability problem, I performed both tests of parametric ANOVA with within-factor Movement Difficulty (tunnel vs. no tunnel) and within factor of Timing (first half vs. second half), on all the performance and usability problem metrics. Because parametric assumptions of normality were not met, I also performed nonparametric Wilcoxon Signed Rank tests on the same factors; whenever the nonparametric and parametric tests agreed on significant findings, I report the parametric results.

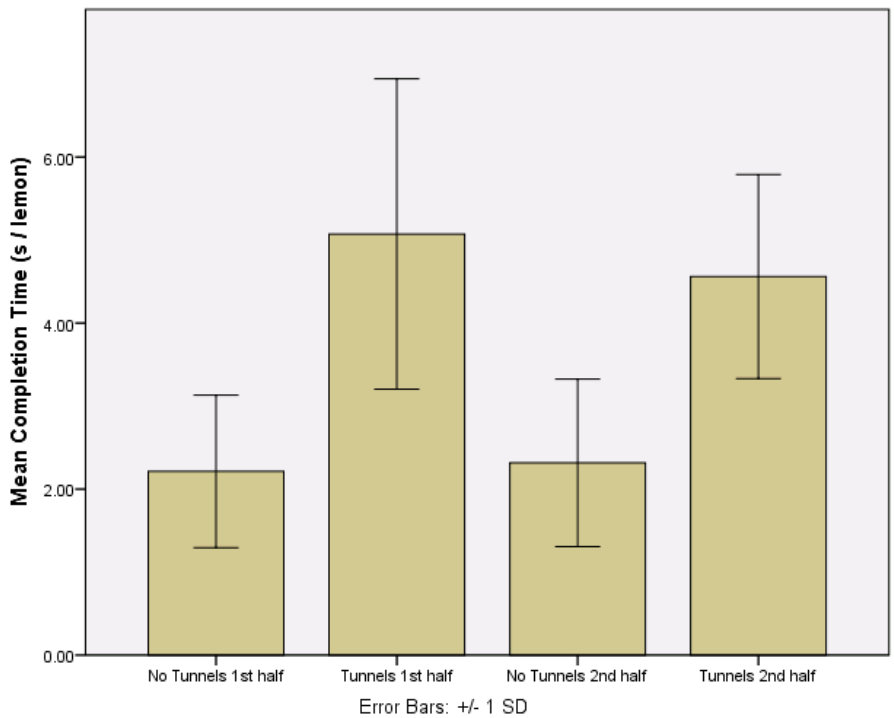


Figure 4.5.17. Average task completion times for each level played in the game. Significant differences exist between Tunnel levels between 1st and 2nd half of the gameplay.

The analysis for Task Completion Time detected a **significant interaction effect between Timing and Movement Difficulty** ($F(1,39)=13.66$, $p=0.001$). Two Bonferroni-corrected one-tailed paired t-tests were performed to determine if performance increases occur in either Tunnel or No Tunnel conditions. **Performance in Tunnel levels was significantly improved between the first** (mean 5.1 s / lemon) and **second half of gameplay** (mean 4.6 s / lemon, a 10% increase) ($p=0.042$). This may occur because children become comfortable moving around the physical space in which they are playing the game, and/or they become better at understanding how the 3D virtual tunnels viewed on the phone relate to the physical environment.

This effect was only present for the variable of task completion time, and not for other variables such as selection errors, number of tracking losses, etc. **No other significant practice effects were found** when investigating differences in other performance and usability problem metrics. When considering the other metrics, this analysis was unable to detect improvements in children's performance or usability problems, as measured by differences between the first and second half of gameplay.

CHAPTER 5

DISCUSSION

5.1 SUMMARY OF RESEARCH QUESTIONS AND RESULTS

This research has been driven by three core research questions, investigating the relationship between children's age, AR design, and usability as measured by performance and usability problems. In this section I will provide a summary of the research findings associated with each research question. A set of design guidelines will then be presented in the next section, followed by a detailed discussion of the findings.

5.1.1 RQ1: How does children's age relate to performance and usability issues in handheld-AR?

Based on the research findings, the relationship between age and usability is complex and depends on which usability metrics are being investigated. Most metrics show a general linear improvement with increase in children's age, although some usability metrics were much weaker in 5-6 year olds than across the 7-10 years range. Some metrics were similar across all age groups, and other metrics showed an inverse trend of increasing frequency in older children.

Most metrics showed usability improvement with age. For example, age was negatively correlated with the number of problems encountered overall by children, number of problems with tracking loss and recovery, and number of problems relating to orienting the body. However, the 5-6 year old children were sometimes significantly different than the 7-10 year olds, while no differences were detected within 7-10 year olds. For example, 5-6 year olds were significantly slower, had more tracking losses while moving, and took longer time to recover tracking, than compared to the 7-10 year olds. Grips used by children showed an increase in the use of crab-style grips and a decrease in curl-style grips as children get older (based on descriptive statistic trends; not statistically significant), and younger children had difficulties performing interactions that required manipulation of multiple objects, such as holding a phone with one hand and an AR paddle in the other (based on informal observations; not statistically tested).

Several usability metrics did not show significant differences between the 5-10 year old age range. Children across all age groups encountered similar a number of selection errors, instances of bumping or tripping, instances of phone dropping, instances of children covering the camera with the finger, and needing initial instructions of how to use the crosshair interaction. There were also no significant age differences in self-reported fun, ease of use, and comfort when measured in the ARC study, indicating that children in all age groups had similar overall enjoyment and comfort while completing the experimental game.

Finally, there were some usability metrics that showed significant detriments with age, as older children were found to encounter more such events. For example, age was significantly correlated with higher number of observed posture strains and scratching behaviors.

5.1.2 RQ2: How do different handheld-AR interaction techniques compare, in terms of performance and usability issues encountered by children?

I have studied several interaction techniques through my early informal studies, as well as in the ARC experiment. The early studies investigated a wide range of interaction techniques in different applications, while the ARC study investigated crosshair vs. finger touch-based selection, under different movement conditions.

Not surprisingly, the informal observations in early studies showed that interactions requiring multiple hand coordination (e.g., holding a device with one hand while the other hand manipulates an AR paddle) are more difficult to perform than simply touching on the screen, and these interactions were very difficult for younger children. However, interactions which required no independent hand movement, such as simply moving the device to a proper distance from an AR marker, also generated problems for children, because children were observed to have difficulty determining the appropriate distance where the device should be positioned.

Two popular interaction techniques were compared in the ARC study. The study systematically compared the interactions of selecting items by touching on the screen with the finger vs. selecting items by aiming a crosshair in the middle of the screen. Across all 5-10 year old children studied, the finger interaction was rated as more easy to use, and showed faster performance than the crosshair interaction. There were significant ergonomic differences, whereby children used curl-style grips more often in the crosshair condition than in the finger condition. There were also developmental differences found in the use of these interactions. Using a finger to select items quickly and accurately was correlated to skills used when children construct structures with toy blocks; however, using a crosshair to do these tasks was correlated to skills involved in hand-eye coordination and fine-motor precision. Interestingly, there were no significant differences detected between interactions in terms of self-reported fun and comfort.

The ARC study also systematically studied the effects of movement conditions. In levels where children had to move more and select items inside of tunnels, children encountered more problems such as selection errors, more issues with losing and recovering tracking, more encounters of posture strains, and more bumping/tripping. Self-reported comfort and ease of use was not statistically different between movement conditions; however, when asked to explicitly choose which levels were more fun, older children rated tunnel levels as being more fun than non-tunnel levels.

5.1.3 RQ3: What types of usability issues are experienced by children in handheld-AR?

Children ages 5-10 years old have been observed to experience a variety of problematic behaviors while using handheld AR. Table 5.1.1 lists the issues which have been identified through the ARC study and through my research on children's AR usability framework (Radu and MacIntyre 2012). The issues have been categorized according to the domain of child development which is likely the main cause of each issue.

Table 5.1.1. Problematic behaviors encountered by children studied in this research. ("CORREL AGE": behaviors are significantly positively or negatively correlated with age. "OBSERVED AGE": observed to be related to age, not statistically tested. "AGE 5-6 > 7-10": statistically higher frequency only in 5-6 year olds. "TUN > OPEN": statistically more frequent in levels involving tunnels. Tests at $p<0.05$)

Issue category	Significant age differences?	Significant interaction differences?
MANIPULATION		
Losing tracking while walking	**AGE 5-6 > 7-10**	**TUN>OPEN**
Losing tracking by covering the camera with the finger	**Not sig.**	**Not sig.**
Strained body posture (e.g., Standing bent on a table)	**CORREL AGE +**	**TUN>OPEN**
Dropping the phone	**Not sig.**	**Not sig.**
Strained grip (e.g., Stretching hand)	**Not sig.**	**Not sig.**
Difficulty holding a phone in one hand, and using other hand to move marker	**OBSERVED AGE -**	NOT TESTED
Difficulty using a marker to intercept a moving virtual object	**OBSERVED AGE -**	NOT TESTED
Difficulty holding the device steady while performing interaction	**OBSERVED AGE -**	NOT TESTED
Difficulty moving a marker on a specified path	NOT TESTED	NOT TESTED
Difficulty turning body around to look at virtual panorama	NOT TESTED	NOT TESTED
SPACE		
Losing tracking by aiming the camera away from the gameboard	**CORREL AGE -**	**Not sig.**
Difficulty orienting body in relation to the gameboard	**CORREL AGE -**	**TUN>OPEN**
Losing tracking by aiming the camera too close to the gameboard	**Not sig.**	**Not sig.**
Not understanding that the virtual world is a 3D space anchored to the physical world, instead of a 2D image on the screen	**OBSERVED AGE -**	NOT TESTED
Difficulty moving the device to the proper distance which initiates interaction	**OBSERVED AGE -**	NOT TESTED
Difficulty moving the device to the orientation which initiates interaction	NOT TESTED	NOT TESTED
Inability to remember the configuration of a large virtual space	NOT TESTED	NOT TESTED
Not detecting when a virtual item is on top of a physical item	NOT TESTED	NOT TESTED
Difficulty predicting what virtual objects are visible from another angle	NOT TESTED	NOT TESTED
ABSTRACT THINKING		
Difficulties interpreting tracking loss and recovering tracking	**CORREL AGE +**	**TUN>OPEN**
Needing initial instruction on how to use crosshair	**Not sig.**	**Not sig.**
Needing in-game instruction on how to use crosshair	**Not sig.**	**Not sig.**
Not understanding the game storyline	**Not sig.**	**Not sig.**
Not understanding general game mechanics	**Not sig.**	**Not sig.**
Not understanding that virtual objects are computer generated, and they do not need to obey physical laws	**OBSERVED AGE -**	NOT TESTED
ATTENTION		
Bumping or tripping	**Not sig.**	**TUN>OPEN**
Interruption due to scratching	**CORREL AGE +**	**Not sig.**
Interruption due to self-distraction	**Not sig.**	**Not sig.**
Understand that virtual objects are computer generated, and they do not need to obey physical laws	NOT TESTED	NOT TESTED

5.2 GUIDELINES FOR DESIGNING USABLE HANDHELD AR FOR ELEMENTARY-SCHOOL CHILDREN

In the following subsections, I provide a set of guidelines primarily intended for designers of handheld augmented reality experiences for children. The first section describes a summary, while the following sections provide guidelines specifically for designing to 5-6 year olds, 7-8 year olds, and 9-10 year olds.

5.2.1 GUIDELINES: GENERAL SUMMARY

- Age influences performance and usability issues encountered
    - Expect children aged 5-10 to show improved AR performance as they become older, although note that some issues increase in frequency as children get older.
    - Expect a significant gap between 5-6 and 7-10 year olds
        - The younger children are significantly slower (at least 30% slower than 7-10 year olds), significantly more prone to losing tracking if the game requires movement (at least 62% more errors), and significantly slower to recover tracking (53% more slower).
        - Because of the 5-6 vs 7-10 split in ages, basic interactions in AR games for 10 year olds should be usable by 7 year olds, but games designed for 7 year olds may be difficult for 5 year olds
    - Children of all ages may drop the phone, bump or trip when moving, and lose tracking by covering the camera with the finger.
    - Some issues occur more frequently in order children, such as more body strain and more bumping/tripping.
- Design for fixed tables and varying child heights:
    - The tables used in children's schools and homes are typically not configurable to different heights. However, children's heights will differ, thus influencing their play angles.
    - Ideally, AR experiences should detect the child's angle to the gameboard, and adapt the game to that. This can be done through smart tutorials.

- Be aware of the way children grip devices
  - Children's grips change over time: young children use lots of curl grips; older children use more crab grips. This is likely due to increased hand size and strength.
  - Some grips are detrimental to performance (e.g., straight grips are correlated to frequent tracking losses). Children should be trained proper grips in the game tutorial.
  - Children will get tired if they have to hold up a device, and/or if they have to move or bend their body around the gameboard. It's important to design experiences where children can take breaks or sit down.
- Children have problems with tracking loss and recovery.
  - Make tracking loss explicit because some children may not notice it, and do everything possible to help kids recover it (e.g., use sensor fusion).
  - All age groups have frequent issues losing tracking while walking or putting fingers in the camera.
- Moving is fun for older children, but problematic for younger children.
  - Requiring kids to move around and change perspective leads to more issues with tracking losses, strained postures, bumping into objects.
  - Younger children especially have problems with moving around the gameboard, and with detecting and recovering tracking loss. Avoid frequent movement and tracking loss opportunities in games with young children.
- Fingers and Crosshair selections are both usable by young children.
  - Both interaction mechanics are similar in terms of selection accuracy, both receive high ratings for fun and comfort across age groups, however Crosshairs are rated as generally less easy to use (by 10%).
  - Crosshair is sometimes better. It can be used with large devices without straining hands or requiring independent hand movements. Trends indicate that this is more comfortable for young children 5-6 years old, and leads to less tracking errors.
  - Finger selection is better than Crosshair selection on time (by average 25%) and ease of use (by 10%). Trends show that finger selection is better on accuracy, and seems more comfortable for 7-10 year olds, generally seems to lead to less tracking losses.
  - Different developmental skills are engaged in different interactions. Finger interactions are correlated to block construction skills; Crosshair interactions are

correlated to hand eye coordination and spatial skills. Children with different skills may show different performance and prefer specific interactions.

  - Like any UI metaphors, AR interaction mechanics are easier if they are familiar. Children are used to touching with the finger on touchscreens; but crosshair instructions are required with most children.

- Previous experience with technology is associated with inaccuracy and more tracking losses. This is potentially due to increased comfort manipulating non-AR games. Tutorials should be designed to emphasize the differences between AR and traditional experiences.

- A good tutorial is important:
  - The tutorial should measure the height and performance of children, so that the game can be adapted to the child's body, age, and physical environment.
  - Children are probably used to playing 2D games on touchscreens, they will need to be taught how this game differs from 2D games
  - In the tutorial emphasize that the game space is 3D, by having children perform tasks that require changes in perspective
  - Encourage children to move around the gameboard so they know their body movements affect the game
  - If touching on screen doesn't apply to the game, explicitly give feedback users learn right away
  - Teach users to avoid touching the camera
  - Teach users different ways of gripping the device
  - Teach users about the tracking technology
    - Specifically show them the conditions when tracking gets lost (e.g., covering the camera, pointing away, too close, too fast)
    - Specifically describe how to recover from tracking loss (e.g., teach the proper position and orientations where tracking is most likely to be regained)

5.2.2 GUIDELINES: DESIGNING FOR 5-6 YEAR OLDS

What is highly difficult for 5-6 year olds?

- High difficulty interacting with a handheld AR experience by holding a smartphone with one hand, while moving a trackable object with the other hand.
  - Tip: Place the AR screen on a stand, so that children's hands are free to manipulate physical props.

- High difficulty dragging or drawing a finger on the smartphone touchscreen, while looking at an AR scene.
  - Tip: Place the AR screen on a stand, so that the device is stable and children don't need to hold it.
  - Tip: Design interactions where children do not need to hold the phone while drawing their finger (e.g., take a photo of a real object, then draw on the photo)

- High difficulty understanding and recovering from tracking loss (all children had this problem, 36% needed the experimenter to physically intervene to solve this problem)
    - Tip: Design games where the chance of tracking loss is minimal.
    - Tip: Do not require the player to move around the gameboard, or perform fast actions by moving the camera, or perform actions moving the camera too close or too far from the gameboard.

- High difficulty orienting the body, either to collect items placed around obstructions, or to move in order to recover tracking (43% of children needed the experimenter to physically intervene to solve this problem)
    - Tip: Design games where young children are not required to move their body around virtual obstructions in the 3D game space.
    - Tip: Walking around a gameboard is correlated to higher frequency of tracking losses. Games requiring player movement should either reduce challenge, so that players can be careful when moving the camera, or be more forgiving to instances of tracking loss.
    - Tip: AR games which require children to move around may cause children to trip or bump into their physical environment. Designers should ensure that children are warned about such potential dangers, and encouraged to clear their surroundings before playing, and be attentive during gameplay.

What is challenging for 5-6 year olds ?

- Challenge performing fast actions (young children are ~30% slower at collecting virtual items than 7-10 year olds), or fast recovery of tracking (young children are ~53% slower at tracking recovery than 7-10 year olds)
    - Tip: Expect 5-6 year old children to perform differently than 7-10 year olds, and design performance-based challenges for their specific level of performance.

- Some challenge holding the phone with controlled force (14% of children dropped the phone)
    - Tip: Due to small hands, expect young children to naturally use “curl” style grips when holding mobile devices, especially if devices are large.
    - Tip: Some children will use a “straight” grip which is correlated to more tracking losses due to finger occluding the camera. Train children to avoid this grip.
    - Tip: Early in the AR experience, show children how to hold the device properly either in “curl” or “crab” grips.
    - Tip: When designing for large mobile devices (e.g., tablets), young children will most likely be unable to hold the device while touching the screen, so design for crosshair-based selection rather than finger-based selection.
    - Tip: Encourage parents to use protective casings for mobile devices.

- Some challenge using crosshair interaction
    - Children needed instructions on how to use the crosshair (71% of children needed instruction how to use crosshair during the tutorial, 14% needed to be reminded during the game)
    - Compared to Finger selection, the Crosshair selection was rated more challenging. However children reported the Crosshair to be just as fun and

comfortable. Also, crosshair requires different grips and different developmental skills, which may appeal to some children more than others.
    - Tip: Use crosshair as an added challenge in games
    - Tip: To decrease the challenge of aiming precisely, consider using aim assistance such as snapping to nearby items.

- Potential challenge due to young children's short height, as the AR gameboard may be placed either on a table or on the floor.
    - If the game is placed on a table and children are standing, they will have a low angle of gameplay. If children are sitting on a chair, they will be unable to move. If the game is on the floor, children may hold an ergonomically difficult position which might not allow movement.
    - Tip: Designers should avoid requiring children to move around the gameboard, and should be aware of the player's perspective during gameplay.

What is easy for 5-6 year olds ?

- Easy to use games where interaction is performed by rotating the device, or moving the device close to virtual items
    - Tip: Design games where young children can hold the device with both hands and are not required to perform complex or precise motions with their hands or body.
    - Tip: When designing distance-based interactions, provide clear feedback about whether the player is too far or too close from an item.

- Accuracy for selecting items was not statistically different than 7-10 year olds (average 1 errors per item selected using finger or crosshair selection).
    - Tip: Young children may be able to play games with the same precision as 7-10 year olds, although more slowly, due to their developing muscle coordination and precision skills.

5.2.3 GUIDELINES: DESIGNING FOR 7-8 YEAR OLDS

What is highly difficult for 7-8 year olds?

- This research did not identify extreme difficulties for 7-8 year olds using handheld AR. It is expected they will be able to use handheld AR applications similar to those involved in this research. Some interactions can be challenging, as described below.

What is challenging for 7-8 year olds?

- Challenge interacting with a handheld AR experience by holding a smartphone with one hand, while moving a trackable object with the other hand.
    - Tip: To ease this challenge, place the AR screen on a stand, so that children's hands are free to manipulate physical props.

- Challenge dragging or drawing a finger on the smartphone touchscreen, while looking at an AR scene.
    - Tip: To ease this challenge, place the AR screen on a stand, so that the device is stable and children don't need to hold it.
    - Tip: Or, design interactions where children do not need to hold the phone while drawing their finger (e.g., take a photo of a real object, then draw on the photo)

- Some challenge understanding and recovering from tracking loss (14% needed the experimenter's verbal help to solve this problem)
    - Tip: Design games where the chance of tracking loss is minimal.
    - Tip: To avoid this issue, do not require the player to move around the gameboard, or perform fast actions by moving the camera, or perform actions moving the camera too close or too far from the gameboard.

- Some challenge orienting the body, either to collect items placed around obstructions, or to move in order to recover tracking (50% of children appeared challenged by needing to reorient themselves, although only 1 child - 7% of total - needed the experimenter to verbally help to solve this problem)
    - Tip: Moving around can be an added challenge which improves engagement and motivation. Games can be designed to require the user to solve tasks while looking and moving around 3D virtual obstructions.
    - Tip: To ease this challenge, decrease the complexity of the 3D environment which children move around, or design games where young children are not required to move their body around virtual obstructions in the 3D game space.
    - Tip: Walking around a gameboard is correlated to higher frequency of tracking losses. Games requiring player movement should either reduce challenge, so that players can be careful when moving the camera, or be more forgiving to instances of tracking loss.
    - Tip: AR games which require children to move around may cause children to trip or bump into their physical environment. Designers should ensure that children are warned about such potential dangers, and encouraged to clear their surroundings before playing, and be attentive during gameplay.

- Some challenge holding the phone with controlled force (1 child, 7% of total, dropped the phone)
    - Tip: Due to small hands, expect these children to naturally use "curl" and "crab" style grips when holding mobile devices.
    - Tip: Early in the AR experience, show children how to hold the device in a controlled grip, either as "crab" or "curl".
    - Tip: Some children will use a "straight" grip which is correlated to more tracking losses due to finger occluding the camera. Train children to avoid this grip.
    - Tip: When designing for large mobile devices (e.g., tablets), young children will most likely be unable to hold the device while touching the screen, so design for crosshair-based selection rather than finger-based selection.
    - Tip: Encourage parents to use protective casings for mobile devices.

- Some challenge using crosshair interaction
    - Children needed instructions on how to use the crosshair (79% of children needed instruction how to use crosshair during the tutorial, 7% needed to be reminded during the game)

  - Compared to Finger selection, the Crosshair selection was rated more challenging. However children reported the Crosshair to be just as fun and comfortable. Also, crosshair requires different grips and different developmental skills, which may appeal to some children more than others.
  - Tip: Use crosshair as an added challenge in games
  - Tip: To decrease the challenge of aiming precisely, consider using aim assistance such as snapping to nearby items.

What is easy for 7-8 year olds ?

- Very easy to use games where interaction is performed by rotating the device, or moving the device close to virtual items
  - Tip: Design games where young children can hold the device with both hands and are not required to perform complex or precise motions with their hands or body.
  - Tip: When designing distance-based interactions, provide clear feedback about whether the player is too far or too close from an item.

- Speed, accuracy, number of tracking losses, and speed to recover tracking were not statistically different between 7-8 and 9-10 year olds.
  - Tip: These children are able to play games with the same performance as 9-10 year olds. However, note that their motivation and usability problems may be different, thus a game designed for 9-10 year olds may not necessarily work as well for 7-8 year olds.

5.2.4 GUIDELINES: DESIGNING FOR 9-10 YEAR OLDS

What is highly difficult for 9-10 year olds?

- This research did not identify extreme difficulties for 9-10 year olds using handheld AR. It is expected they will be able to use handheld AR applications similar to those involved in this research. Some interactions can be challenging, as described below.

What is challenging for 9-10 year olds?

- Some challenge holding the phone with controlled force (17% of children dropped the phone)
  - Tip: Expect these children to naturally use “crab” style and “curl” style (to a lesser degree) when holding mobile devices.
  - Tip: Early in the AR experience, show children how to hold the device in a controlled grip, either as “crab” or “curl”.
  - Tip: Some children will use a “straight” grip which is correlated to more tracking losses due to finger occluding the camera. Train children to avoid this grip.
  - Tip: When designing for large mobile devices (e.g., tablets), children may be unable to hold the device while touching the screen, so design for crosshair-based selection rather than finger-based selection.
  - Tip: Encourage parents to use protective casings for mobile devices.

- Possibly some challenge using crosshair interaction
    - Compared to Finger selection, the Crosshair selection was rated more challenging. However children reported the Crosshair to be just as fun and comfortable. Also, crosshair requires different grips and different developmental skills, which may appeal to some children more than others.
    - Children needed instructions on how to use the crosshair (58% of children needed instruction how to use crosshair during the tutorial, 0% needed to be reminded during the game)
    - Tip: Use crosshair as an added challenge in games
    - Tip: To decrease the challenge of aiming precisely, consider using aim assistance such as snapping to nearby items.

- Possibly some challenge understanding and recovering from tracking loss (16% of children appeared frustrated but were able to resolve the issue by themselves)
    - Tip: Design games where the chance of tracking loss is minimal.
    - Tip: To avoid this issue, do not require the player to move around the gameboard, or perform fast actions by moving the camera, or perform actions moving the camera too close or too far from the gameboard.

- Possibly some challenge dragging or drawing a finger on the smartphone touchscreen, while looking at an AR scene.
    - Tip: To ease this challenge, place the AR screen on a stand, so that the device is stable and children don't need to hold it.
    - Tip: Or, design interactions where children do not need to hold the phone while drawing their finger (e.g., take a photo of a real object, then draw on the photo)

- Possibly some challenge interacting with a handheld AR experience by holding a smartphone with one hand, while moving a trackable object with the other hand.
    - Tip: Older children can easily hold a smartphone with one hand, while touching the screen with the other. It is unclear if they can hold a tablet device with one hand while having the other hand free.
    - Tip: Older children may be able to use a paddle for controlling virtual items, but the game needs to provide clear depth cues and feedback to guide the interaction.

What is easy for 9-10 year olds ?

- Easy and reportedly more fun to use games requiring orienting the body; however, posture strains increase due to children's increased height (41% of older children showed poor ergonomic postures)
    - Tip: Older children preferred game levels that required them to be looking around virtual 3D obstacles. Moving around can be an added challenge which improves engagement and motivation. Games can be designed to require the user to solve tasks while looking and moving around 3D virtual obstructions.
    - Tip: Children may bend their body when looking down at the gameboard, or when looking around obstacles. Design games that account for children's height, and that encourage children to move around, in order to vary their body postures.
    - Tip: Walking around a gameboard is correlated to higher frequency of tracking losses. Games requiring player movement should either reduce challenge, so that

players can be careful when moving the camera, or be more forgiving to instances of tracking loss.
    - Tip: AR games which require children to move around may cause children to trip or bump into their physical environment. Designers should ensure that children are warned about such potential dangers, and encouraged to clear their surroundings before playing, and be attentive during gameplay.

- Very easy to use games where interaction is performed by rotating the device, or moving the device close to virtual items
    - Tip: Design games where young children can hold the device with both hands and are not required to perform complex or precise motions with their hands or body.
    - Tip: When designing distance-based interactions, provide clear feedback about whether the player is too far or too close from an item.

- Speed, accuracy, number of tracking losses, and speed to recover tracking were not statistically different between 7-8 and 9-10 year olds.
    - Tip: These children are able to play games with the same performance as 7-8 year olds. However, note that their motivation and usability problems may be different, thus a game designed for 7-8 year olds may be too easy and not necessarily as engaging for 9-10 year olds.

5.3 DETAILED DISCUSSION

5.3.1 Behaviors that do not vary between age groups

There were no statistically significant differences between all age groups on number of selection errors, instances of bumping or tripping, instances of phone dropping, instances of children covering the camera with the finger, and needing initial instructions of how to use the crosshair interaction. There were also no significant age differences in self-reported fun, ease of use, and comfort.

Bumping or tripping behaviors occurred in roughly 37% of children across each age group. Children were sometimes observed bumping their feet into the legs of the table, or trip when walking around the gameboard. This inattentive behavior has been reported in many reports of mobile AR applications, even with adults, and it is likely an indication of the player being highly focused on the game. Handheld AR designers need to inform the user to be aware of hazards in their environment while playing with the applications.

Covering the camera with the finger was encountered by roughly 70% of children across age groups, and the median number of occurrences per child decreases as children become older. The camera in the experimental study was located on the top left of the Atrix HD smartphone. Covering the camera with a finger is thus related to the grip which children use on their left hand, and a significant correlation was found to the use of a “straight” grip. Designers should teach children to know where the camera is located, and to use proper grips which avoid covering the camera. Another reason for covering the camera with a finger may be that children become tired and switch their grip more often; it is unclear if this is a significant reason, as this metric was not analyzed in the present research. A third reason may be that children become very focused on their actions in the game, and fail to notice when their finger starts to occlude the camera; this may occur more often in younger children, as they are known to have a limited focus of attention. Finally, there was a significant relationship found between previous technology experience, as children with more previous experience with technology were seen to cover the camera with the finger more often. Such children may be comfortable holding devices in a way that is detrimental to camera-based AR applications, and designers should emphasize the fact that AR is not like previous technologies which children may be comfortable with, thus they should pay attention to how they grip the device..

Children across all age groups required instructions on crosshair use, with a non-significant drop in older children. This indicates that children are familiar with finger-based interactions, and not so familiar with crosshair-based interactions. Thus designers should train children how to use crosshairs, and should clearly highlight when the game interaction mechanic changes to require crosshair-style input.

Dropping the phone did not show statistical differences between age groups, although this event was not encountered in many children, and when encountered the median occurrence was once per affected child. This occurs due to children’s developing grips, muscles, and multi-tasking capabilities. Designers should ensure that children are trained to use the proper grip, and should encourage parents to supply protective casings for children’s AR-enabled devices.

There were no statistical differences found between age groups on the metric of number of selection errors. Although there is a general increased trend in lower accuracy with age, the 7-8 year olds group showed a higher variation and average number of selection errors on Tunnel levels than compared to the other groups. It's unclear why this effect happens, but one reason may be related to grip styles or increased body height: children 7-8 years old start to use "crab" style grips more than younger children, and they start encountering body posture strains; this combination may make children more uncomfortable and thus impatient. It is also worth noting that this effect of 7-8 year olds having higher number of selection errors was not statistically significant, thus it may be simply a sampling artifact.

Other metrics that did not show a trend across age is self-reported fun, ease of use, and comfort. These remain quite constant across the age groups (mean ratings for fun is 4.3/5, ease of use 4.2/5, comfort 4.1/5). This is a positive finding for informing AR design, indicating that children across the 5-10 age group can play handheld AR games and are willing to. Especially positive is the fact that self-report ratings are high even in 5-6 year olds, who experienced a higher amount of issues while playing the game compared to the other age groups. This indicates that young children are able and willing to engage with handheld AR games that use finger and crosshair interactions.

5.3.2 Usability improvements around 7 years

Generally, there is a visible trend whereby children's performance increases with age. However, in the 5-10 year old range there appear to be two significantly different groups of children: the 5-6 year olds, and the 7-10 year olds. Specifically, this study shows that under a variety of conditions, young children aged 5-6 years old are significantly different than 7-8 and 9-10 year old groups, while statistical analysis does not find differences in the 7-8 and 9-10 year old groups. This occurred over multiple metrics such as task completion time, number of tracking losses, time to recover tracking, and usability issues related to losing tracking while walking. In all these cases, it appears that a developmental threshold exists whereby children 7-10 are similar to each other, while children 5-6 are different. To developers of augmented reality this can signify that this age range should be treated as two different age groups, and caution should be exercised when developing for the 5-6 year olds because they will exhibit low performance and usability issues: 5-6 year old children will perform much slower, encounter much more tracking losses, take longer time to recover tracking, will have trouble losing tracking while walking, and will have trouble orienting themselves around the gameboard.

A noteworthy finding is that, even though the 5-6 age group encountered more errors than the other groups, they also reported the highest levels of fun (average 4.5 / 5, not significantly different than other age groups) and self-reported "ease of use" consistent with the other age groups (average 4.2/5, not significantly different than other age groups). This research shows that children as young as 5 years old are able to engage with smartphone-based AR games which require a high amount of physical manipulation and spatial reorientation, even though these experiences lead them to have higher amount of issues than other age groups. I believe this perseverance and high ratings of self-report are due to motivation, specifically influenced by the fact that the game storyline and end-of-level game sequences were specifically designed for (and strongly informed by) younger children, and the motivation is also influenced by the introduction tutorial which taught children how to hold the phone and move around the augmented space. The development of the game storyline, end-of-level game sequences, and tutorial, was done through frequent interactions with young children during design and playtesting phases. For AR game designers, this suggests that involving children early in the design process is extremely valuable for creating a motivational game, which children will persevere through even though they may encounter higher usability issues.

5.3.3 Usability improvements across all age groups

There are some general correlations and trends observed throughout the spectrum of 5-10 year olds showing that as children become older, their ability to use AR applications improves. For example, significant correlations were found between children's age and the overall number of usability issues, notably for higher severity issues; also correlations between age and issues related to tracking losses due to aiming away, and issues related to interpreting and recovering tracking. In all these cases, children encounter less issues as they get older. Finally, there are trends observed in the descriptive statistics, whereby older children encounter less issues related to finger occlusion, they report more comfort in tunnel levels, they show lower number of issues related to understanding how to use crosshair interactions, and lower number of issues related to understanding game mechanics. Observations from pilot studies also indicate that younger children have more difficulties than older children with complex interactions, such as manipulating paddles while holding a phone, or holding the phone while dragging and dropping on the screen.

Within the 7-10 age range there were no statistically significant differences found. However, there are interesting non-statistically-significant observations in relation to number of tracking losses and grips used by these two age groups. The statistical analysis showed that, while the mean number of tracking losses decreases with age across all groups, for 7-8 year olds the difference in tracking losses between Tunnel vs No Tunnel conditions is not statistically significant, while for 9-10 year olds this difference is statistically significant (similarly for 5-6 year olds). The use of grips also appears different between these ages (although no statistical effects of age were found), as indicated in Chapter 4. Notably, children in this age group tended to use more straight grips, which were shown to be significantly correlated to tracking losses due to fingers in front of the camera. These observations indicate that there may be some differences in performance between 7-8 and 9-10 year old groups, and more focused research is required to investigate those differences.

5.3.4 Problematic behaviors increasing with age

Interestingly, as children become older they may encounter new kinds of problematic behaviors. As children get older, they show significantly more body strains and interrupt the gameplay significantly more often in order to scratch. They also show (statistically non-significant) increasing trends in bumping and tripping, and trend to decrease in having fun. The issues related to body strains, and bumping and tripping can be problematic. They indicate that, although older children may not report poor comfort levels, they are more likely to hold poor ergonomic postures while playing games (e.g., straining their back or neck) and perform quick and potentially dangerous movements (e.g., quickly jump around to change perspective). These may be problematic if the child is playing a game for a long time and/or in a busy outdoor environment. In order to ameliorate these issues, games designers should design games where the natural posture is not uncomfortable for tall children (e.g., play games on a vertical poster rather than on a low table), encourage children to move around (e.g., change body postures in order to perform game tasks), and not encourage rapid movements (e.g., slow the gameplay in order to avoid tripping) when designing for outdoors or physically cluttered environments. The latter two behaviors that are directly related to age, namely interrupting by scratching and having lower amounts of fun, may not be general problems, and instead may be indicative that the ARC game was too easy for the older group. The game was designed for younger children, thus it is foreseeable that if a game is made entertaining and challenging enough, older children will be more focused and have more fun (for a discussion on this topic, see Chapter 5.4.1).

5.3.5 Finger vs. Crosshair Selection: Differences in Usability and Developmental Skills

Overall, the data indicates that the Finger selection was easier to use than the Crosshair selection interaction type. When using the Finger selection interaction, children completed the task significantly faster, and they reported significantly higher levels of ease-of-use than compared to the Crosshair interaction. However, the developmental test measures indicate that different cognitive- and physical-skills are underlying these two interaction types. Users behave differently when an AR interfaces is designed with different kinds of selection type, and this may be for several reasons. The type of interaction can cause the user to grip the device in different ways (e.g., a “curl” grip is more common in crosshair interactions than in finger interactions), and this can require the use of different physical manipulation skills. The type of interaction can also add an extra layer of indirection between the user and the game (e.g., to select items in the finger condition, a player merely needs to touch an item on the screen; but in the crosshair condition, the player first needs to aim the device at the target, then touch on the screen), and this might increase the user’s cognitive load as they use the interaction. Finally, one interaction type may be more difficult or require more precision, and this may cause users to use a high degree of attention in order to achieve the interaction. These differences cause users to use different physical and cognitive skills when using specific interactions such as finger or crosshair selection.

The Finger selection interaction requires children to hold the device steady primarily with one hand, while the other hand is moved such that the finger touches the screen at the appropriate target location. The analysis of children’s AR performance in relation to their developmental test scores, indicated that Finger Selection levels were significantly correlated to block-construction skills in both time to complete and number of selection errors, during tunnel conditions. This data indicates that this kind of interaction relies more on the user's ability to manipulate objects with their hands.

In contrast, the Crosshair selection interaction requires children to aim the crosshair by precisely reorienting the device (likely by coordinating both hands) such that the center of the screen is aimed at a target; selection is indicated by touching one of the buttons at the side of the screen. Furthermore, in order to reorient the device toward the target, the user must have an understanding of the spatial relationships between the device and the physical gameboard, and must adapt the movement of the device as they are aiming at the item. The fact that children are slower at completing tasks using the crosshair selection could be due to the fact that both hands must be coordinated together in order to aim the device at a target, or that moving the device is generally slower than flicking a finger at a target as with the other selection technique, or that the overall complexity of the interaction is too high and thus causing an increased cognitive load. The analysis of children's developmental tests in the ARC study showed that this selection technique is related largely to visuomotor skills, and also to block construction and spatial relations skills. Selection accuracy using crosshair in no-tunnel levels was correlated to visuomotor precision, indicating that this skill is required for properly aiming at targets using this selection technique. This correlation was not found for crosshair selection errors in tunnel levels; this is possibly because selecting with crosshair in tunnel levels may rely on other skills such as aiming by physically moving one's upper body around the gameboard. Visuomotor precision is, however, related to children's number of tracking losses in tunnel levels using crosshair, and time to complete the task in no-tunnel levels using crosshair. This is possibly because hand-eye coordination is required to reorient the device under a crosshair grip, if it begins aiming away from the gameboard or if it moves away from a target. Overall, in contrast to the finger selection, the crosshair selection relies on visuomotor skills and spatial relations skills, and less on block construction skills. No other significant correlations were found with the 2D spatial relations test besides the relationship to time to recover tracking under the crosshair tunnel condition; it's unclear why, but possibly because the test used was focused on 2D spatial rotations and may not be applicable to the 3D space visible in augmented reality. During the informal observations of participants being trained in the Crosshair selection conditions, I have observed children pointing the screen away from the target object instead of towards it, potentially due to the fact that they did not appropriately understand the spatial relationship between the gameboard and the device. After exposure to the training level, this ceased to be a significant effect; thus, proper instruction is important.

The form factor of the device is another issue to consider when designing AR interaction techniques. In the ARC experiment I used a smartphone, which children could hold with both hands or a single hand. The finger selection interaction is preferable in this case, since children find it faster and more easy to use, and since the phone is light enough to be held with one hand. However, finger-based interaction can become problematic in AR applications for larger devices, such as tablets or larger smartphones. A device may be too large or too heavy for a child to hold while one hand is touching on the screen. In such a case, the Crosshair-based interaction is more suitable as it allows children to grasp the device with both hands while interacting with the application. In my other pilot studies where children used different kinds of devices and interactions, I observed that children had problems holding a smartphone steady while dragging & dropping on the screen and when using other objects like paddles; additionally, children showed difficulties holding up a tablet device and attempting to touch the screen. Device ergonomics issues will be discussed further in Chapter 5.4.12.

5.3.6 Dealing with Tracking Technology

This augmented reality games used in this research were based on Vuforia technology, which uses the phone's camera to track a paper-based printed image (the "gameboard"). The 3D game appears on the printed image when parts of the image are visible through the phone's camera, and it disappears when the image is no longer within the camera view. For the ARC experiment, the smartphone was an Atrix HD with the camera placed on the top-left of the back side of the phone, as is the case with most smartphones.

Tracking loss for the study participants occurred whenever the children would point the camera away from the printed image (either while standing still or moving around the gameboard), or when they covered the camera with their fingers. In the experimental condition where there were No Tunnels, children typically did not change perspective, therefore tracking losses were not caused by movement around the gameboard. The data indicates that the number of tracking losses was significantly inversely-correlated with children's scores on block construction skills (when using finger selection techniques) and visuomotor skills (when using crosshair selection techniques). In order to avoid tracking losses, children need to be aware of where the gameboard is in relation to the device, need to be aware of how much of the gameboard is visible through the phone, need short term memory and abstract understanding of the fact that the camera needs to keep tracking the paper, and they need good physical manipulation and hand-eye coordination skills to recover in case the phone is moving away from the board or if their finger is moving in the way of the camera.

When playing Tunnel levels, which required walking and changes in perspective, children had a significantly higher number of tracking losses in general. Walking and changing perspective created more opportunities for children to cause tracking loss by moving the camera too fast, or aiming the device away from the gameboard, or putting their finger in the way of the camera. The research analysis detected stronger correlations between developmental tests and the number of tracking losses in Tunnel levels, possibly because there were simply more tracking losses in these conditions, or possibly because children were required to use different physical and cognitive skills while reorienting themselves around tunnels.

In order to recover AR tracking, children needed to point the phone camera at the printed gameboard image. The time it took children to recover from tracking loss was significantly different between the 5-6 and 7-10 year olds, but not significantly different between any of the game levels, thus it is not strongly influenced by the style of selection (Finger or Crosshair) or by the movement difficulty (Tunnel or No Tunnel). Recovering tracking appears to be a general process independent of the type of interaction in the AR application. The type of AR interaction, and possibly the grip and posture encouraged by the interaction, do appear to influence the number of tracking losses (as discussed above), but they not influence children's ability to recover tracking.

AR experiences can be designed to minimize the number of tracking losses experienced by players. In this AR game, losing tracking paused the gameplay, thus did not create any negative effect on the child's gameplay. However, I did notice that children became frustrated if they lost tracking frequently. Thus, it is preferable if the AR experience implements features to avoid tracking loss. The AR application can be designed such that players are encouraged to be looking at the gameboard while moving (for example, if the AR application depicts a phenomenon that is interesting while being watched from changing perspectives, like a virtual prism). Furthermore, the AR technology can detect how much of the printed image is visible within the camera, therefore it can display a warning if the child is playing too close to the border of the gameboard, or if the child's finger is starting to occlude the camera.

Finally, the analysis also detected a general correlation between children's number of tracking losses and the use of "straight" grip in the left hand. The camera on the experimental smartphone was placed on the top of the left edge of the device, thus children who used the straight grip were more likely to lose tracking by obstructing the camera view with their fingers. This indicates that grip should be considered an important factor when considering AR experiences. Interactions should be designed such that children's small hands can comfortably perform the interactions. Furthermore, the application should train users where the camera is located, show what happens when the camera is occluded by the finger, and show which grips are comfortable while leading to decreased usability issues.

5.3.7 Tunnels vs. No Tunnels: Perspective, Accuracy and Occlusion

In levels involving No Tunnels, the game created three-dimensional spheres (the lemons) on the gameboard. From the player's default perspective in front of the gameboard, these targets were always visible and not occluded by any other game structures. In levels involving Tunnels, the lemons were encased in three-dimensional tunnels, which required participants to change perspective in order to see the lemons inside the tunnels. Compared to levels with No Tunnels, the levels involving Tunnels led to significantly longer task completion time, more selection errors and higher number of tracking losses. These kinds of effects are expected to occur in other AR experiences where players are required to change perspective around the gameboard.

In order to collect the lemons inside the tunnels, most participants changed their perspective by walking around the gameboard, and some participants bent their body while standing still. Longer task completion times in the Tunnel levels were likely caused by the fact that players had to reorient their body in order to aim at the lemons. It is possible that for some children the task completion times became longer because of difficulties perceiving the virtual game world as a 3D space in which to move the human body, such as may be caused by undeveloped spatial skills or proprioception skills. It is also possible that understanding 3D tunnels around which the player has to move their body, may require high cognitive load and thus reduce children's ability to coordinate their actions quickly.

The three-dimensional structure of the tunnels was also a factor which created issues for selection accuracy. The optimal angle to collect a lemon is to look at it while being aligned with the entrance of the tunnel – this way it appears as a full sphere on the smartphone screen; however, if a player did not look from the entrance of the tunnel, the tunnel walls could occlude the lemons within, thus yielding a smaller selection area. Lower accuracy rates related to Tunnel levels might also be caused by the fact that, when children's bodies are bent, it becomes more difficult to aim at a target, as well as perform the proper physical action required to select the target. Furthermore, as discussed in the previous section, it is likely that moving around the gameboard also contributed highly to the number of tracking losses encountered by children in the Tunnel conditions, because tracking would be lost for various reasons as children moved.

These effects will occur in any AR experience where players are required to change their perspective around three-dimensional content. The higher degree of inaccuracy in these experiences can be a positive factor since it adds challenge to the experience (for a related discussion, see Chapter 5.4.1). However, if desired there are methods for making the experience easier, such as by placing guidance arrows to indicate how the player should change their perspective in order to interact with game items, or by changing the interaction technique such that targeting is more automatic once a part of the target is visible.

5.3.8 Rehabilitation and Skill Learning

In order to interact with an augmented reality experience, children are required to employ different developmental skills. This analysis detected that the visuomotor precision, block construction, and 2D spatial relationships tests were correlated to different kinds of AR interactions, whereby higher ability scores correlated to better performance. Further research needs to investigate if, through repeated exposure to augmented reality experiences, children can further develop these solicited skills. If that is the case, augmented reality interactions may be designed to measure and adapt the challenge level to children who are lacking in specific skills (see Chapter 5.4.2 for a related discussion).

5.3.9 Detrimental Effects of Previous Technology Exposure

Inverse significant correlations were found between (age-standardized) AR performance and their previous experience with technology. More previous experience with technology (such as computers, smartphones, video game consoles and handheld controllers) was correlated with more selection inaccuracies and more tracking losses. There was no correlation between technology experience and completion time, indicating that children with more technology experience did not complete the task significantly faster, even though they had more inaccurate attempts in their interactions. One reason why this might occur is that children who are exposed to more technology may have more knowledge about the robustness of technology, and may feel more comfortable manipulating devices in ways they are familiar with. This may lead them to be more aggressive when interacting with the game items or when moving around the gameboard because they may be used to technology experiences that are forgiving; however, these behaviors will lead to more inaccuracies and more tracking losses. On the other hand, children who have less exposure to technology may be more careful when handling devices or new technology, such as AR-enabled games, thus leading to less incorrect selection attempts and more attention to AR tracking. Finally, previous research has also indicated a relationship between socioeconomic status and exposure to technology and material resources (see Chapter 5.4.7), thus it may be that children with lower exposure to technology are also more careful because they are more fearful of damaging the technology. In the study I have also detected a relationships between gender and technology exposure (see Chapter 5.3.10) and it is unclear how much of the results are due to gender vs. technology exposure (a limitation discussed in Chapter 5.4.4). Finally, there are several limitations to the metric of technology exposure, thus further research is necessary to investigate these effects (for further limitation see Chapter 5.4.9).

5.3.10 Gender Differences

Significant gender differences were found in relation to technology exposure. Boys had on average 19% more technology exposure than girls. Technology exposure was found to mediate gender performance in AR. Without accounting for technology exposure, gender differences were exhibited in AR performance. When accounting for technology exposure, there were no significant effects found between gender and any of the AR performance metrics or usability problems. This indicates that the gender effects are mediated by technology experience, and AR studies investigating gender effects must account for any effects that may be caused by previous technology experience. However, the quality of technology exposure may be different in different genders (Jackson, Zhao et al. 2008), a feature which is not measured in the technology exposure metric used in this study (for further discussion of limitations see Chapter 5.4.4, and Chapter 5.4.9).

5.3.11 No Effect of Hand Size

As children get older, their hands also grow, and size may cause children to more easily control the device. However, as children grow there are also improvements in their physical and cognitive skills, thus it is important to analyze the effects of hand size after accounting for the general effects of age. After accounting for general effects of children's age, hand size was not found to be correlated to AR performance. From our analysis, children with relatively larger hands did not show relatively better AR performance. This result can indicate that either (1) the experimental sample of children did not have high variation in hand size for detecting an effect outside of the general effects of age; or (2) hand size does not lead to strong differences in AR performance.

5.3.12 The Effect of Practice

In this research I have also analyzed differences between child performance and usability problems, between the first and second half of gameplay in the ARC study. Task completion time in Tunnel levels showed improvement, indicating that within the short exposure to the game, children quickly improve their speed of moving around the AR play space. During the initial exposure to the game, it is likely that children are learning multiple things. Children are learning what to expect from the game in terms of game mechanics and challenge, they are learning how to hold the device, they are learning the layout of the physical space as well as how to move around the game space, they are learning how to avoid/recover tracking loss, they are learning how to interpret the 3D virtual content viewed on the phone, they are becoming comfortable being observed by the experimenter, etc. Becoming comfortable with these factors can account for children's improvements in task completion times between first and second half of gameplay. Furthermore, the learning effects in terms of time will become more easy to detect statistically in tunnel levels more than in non-tunnel levels, because tunnel levels generally take over twice as long to complete than non-tunnel levels.

Interestingly, besides task completion time, no other usability metrics showed improvement between first and second half of gameplay. For example, there was no detectable improvement in selection accuracy, number of tracking losses, inability to orient or recover tracking, etc. These behaviors are likely more difficult to learn, and may take more practice to improve. Chapter 5.4.3 will discuss future work in studying longer term practice effects.

5.4 FUTURE WORK AND LIMITATIONS

5.4.1 Leverage Usability Difficulties to Designing for Engagement

Future work should apply the current usability findings to designing engaging educational AR games for young children. The current research has been mainly focused on understanding the usability of handheld AR applications for children 5-10 years old. The research questions specifically investigated usability in terms of performance and usability problems encountered by children, as they used basic AR interaction designs at different ages. From the research I have identified which behaviors children are capable of doing at specific ages, and the degree of difficulty encountered as children perform these behaviors. This knowledge can be used by designers to create AR applications that are suitable for children of a specific age.

However, besides usability, the additional factor of “motivation” plays a critical role in the effectiveness of designing educational applications for children (Malone and Lepper 1987). Educational applications for children typically take the form of games that engage the student in fun activities which have learning outcomes. Motivation is a mediating factor in how strongly a student will engage with an educational application: simply because a game or application is easy to use, this does not mean that children will choose to use it; in fact, players choose to engage with games especially because the game interactions are challenging.

Good games should be easy to learn, and hard to master (Bekker, Barendregt et al. 2005). This research has identified the degree of difficulty which children experience, while using specific handheld AR interactions. Designers can use this knowledge to determine what game elements to incorporate in order to provide the right level of challenge such that players remain motivated. Flow theory (Csikszentmihalyi 1997) models the relationship between personal skill, external challenge, and intrinsic motivation. Applied to game design, it is suggested that good games can keep players highly engaged by maintaining a dynamic balance between the player’s abilities and the challenge offered by the game (Swink 2009, Schell 2014).

Knowledge about the degree of usability of AR interactions can be helpful in designing engaging games and applications for children. Interactions that are known to be too difficult to perform by a certain age group will likely lead to frustration and non-engagement, and should be avoided or potentially designed to be assisted by the game. (For example, requiring 5-6 year old children to move around a gameboard and aim at occluded virtual items is very challenging; this should be avoided, or, a game can allow players to see behind occlusions, or provide a sticky targeting mechanism where players do not need to perform precise orientations). On the other hand, interactions that are too easy to perform should also be avoided because they will likely lead to boredom, unless they are designed to be more challenging. (For example, requiring 9-10 year old children to select targets with a finger on the touchscreen without moving around is an easy task; however, a game can possibly make this task challenging by making the virtual targets change position quickly, or providing multiple options and requiring the player to prioritize or avoid certain targets). Interactions that are known to be somewhat challenging from a usability perspective can be used as game challenges, especially when they are parametrized to be suitable for the specific age group. (For example, children are known to be challenged by aiming at items using crosshair interactions; thus, selecting items with the crosshair will provide some challenge in its basic form, but different age groups will perform the task at different speeds, thus the game needs to provide timed challenges at the speed appropriate for each age group).

In providing age-appropriate design guidelines (Chapter 5.2), I have listed the usability issues as well as the degree of challenge which they pose to children of specific ages. I hope that future work investigates how these varied usability "challenges" can be used to inform the design of games and applications that are appropriately-challenging and engaging for 5-10 year olds.

5.4.2 Design Intelligent Tutorials and Adaptive Games

Future research should investigate how to design effective AR tutorials which not only teach children the proper components of how to use AR experiences, but also tutorials that measure the user in order to adapt the game to the user's characteristics. The tutorial is a powerful component of designing AR experiences for children. When my previous games did not have tutorial components, I observed children experiencing issues such as not understanding that the game is a 3D experience anchored in physical space, not understanding that they can move around the physical space in order to interact from different angles, not understanding why AR tracking is lost, and not knowing how to hold the phone comfortably such that the camera doesn't get covered. During the ARC study I performed several iterations, working with children to design an effective tutorial. In the final version of the experimental game, the tutorial contained several important components which I believe helped children bypass the issues identified above. The tutorial contained phases such as: showing the basic gameplay in a non-AR environment, teaching the child about the phone camera placement and what happens when it is covered, showing a variety of grips, making the child move close/far and around the gameboard in order to see the 3D space, and teaching the child how AR tracking is lost and regained. I believe these components helped to avoid some usability issues, thus I feel the tutorial design greatly influences the issues encountered by children during the game. Future research should investigate the design of effective AR tutorials specifically accounting for the fact that young children may not be able to read or understand complex concepts, that children may not have a parent present when playing such games, and that children's environments may widely differ between uses – for instance having a physically limited play space, or having the gameboard placed on a table or on the floor.

It is also possible to design AR tutorials and games that measure, and potentially adapt to, the user's characteristics and environment. Children will have different heights, and this will affect the typical vertical angle from which children play the game. Furthermore, children may place the gameboard on surfaces of different heights – for example, on the floor, or on chairs, or tables or countertops – and these different heights will also influence the angle and posture from which children play. Additionally, children's physical environments may be limited. For example, children may place the gameboard on a countertop which does not allow them to move to the opposite side of the gameboard, or in a crowded room where they can't move around too much. These environmental properties can be measured during the tutorial level. For example, a tutorial can be designed where the player must try to photograph a virtual character as it moves around a game environment, allowing the game to measure the player's play angle, as well as their ability to move around the physical space, and these measurements can be used to adapt the rest of the gameplay.

The player's developmental skills can also be measured through their performance on AR games. In the ARC experimental game, player's AR performance was correlated to Block Construction, Visuomotor Precision and Spatial Relations skills, thus performance can be used to predict these scores. It may be possible to design AR games that measure different kinds of developmental skills. Further research needs to be done on the reliability of using AR as an instrument to measure child developmental skills. Additionally, it may be possible to design AR games that improve different developmental skills. From the current research, the analysis shows only a correlation between performance and developmental skills, thus it is unclear if the skills can be improved through gameplay. Further research is needed to determine which children's skills can be improved through practice in AR games.

5.4.3 Study Long Term Exposure and Practice

Further research is needed to study how children learn to use AR interfaces over longer periods of time. My studies were short-term exposures to augmented reality interaction techniques. These studies observed how easy it is for children to "pick up and use" the interactions. In the ARC study I analyzed differences between child performance and usability problems, between the first and second half of gameplay. The only metric that showed improvement is task completion time in Tunnel levels, indicating that, within the short exposure to the game, children quickly improve their speed of moving around the AR gamespace. This may occur because children become comfortable moving around the physical space in which they are playing the game, and/or they become better at understanding how the 3D virtual tunnels viewed on the phone relate to the physical environment. However, no other metrics showed improvement.

The limitations of this short exposure are that, if a child had difficulty performing an interaction within this study, this does not mean that the child cannot become fluent at the game over a longer-term exposure with a game they are motivated to play. Future research should investigate how children's performance and usability problems correlate with longer-term exposures to AR technology.

5.4.4 Investigate the Interactions Between Gender and Technology Exposure

The ARC study found significant gender effects in relation to technology exposure, specifically that boys had significantly more technology exposure than girls. I also found that boys showed significantly more selection errors than girls; however, when accounting for technology experience, the gender effect on selection errors disappeared. This indicates that the gender effects are mediated by technology experience, and AR studies investigating gender effects must account for any effects that may be caused by previous technology experience.

However, in the current study I was not able to properly investigate the effects of technology experience irrespective of gender. When not accounting for gender, technology experience was significantly correlated to metrics such as selection errors and number of tracking losses; however, when I removed the effects of gender, the correlation to tracking losses remained but the correlation to selection errors became marginally non-significant. This indicates that gender and technology experience strongly overlap, and a more controlled research study is required to determine the relationship between these variables and their effect on AR performance.

5.4.5 Improve Controlled Experimental Tasks for Studying Children and AR Interaction Techniques

Future research should investigate how to improve the experimental reliability and control, when performing research studies for young children's performance in using AR. Performing interaction technique studies with children in augmented reality has posed specific challenges in the current research. Studies on non-AR touchscreen interactions have previously been done as highly controlled tasks, where each trial presents children with on-screen targets that are simple geometric shapes of constant size. In such 2D applications, the experimenter can control the on-screen target size and location, and can also require an initial finger position in relation to the target. This allows experimenters to standardize the results between participants, and build Fitts' Law models of movement mechanics. In contrast, in handheld augmented reality, a target is a 3D virtual object that is anchored to a physical object (the gameboard) instead of on the screen, and the screen is freely moveable by the user. Therefore, the target's on-screen size and position change in response to the child's distance and orientation to the gameboard; thus the size and position of targets cannot be directly controlled unless the child's movement is controlled. These factors create differences between the item collection trials of one child, and within- and between- children, especially in levels where large amount of movement is required (such as in this levels involving tunnels). This likely leads to less statistical power to detect differences between experimental conditions, and creates difficulties in performing more precise studies such as Fitts' Law investigations.

In the ARC study, I thought about controlling these factors. For example, in order to ensure that all children collect items of the same on-screen size, it is possible to enforce that the game items be a specific constant pixel size on the screen, regardless of the player's distance toward the item. However, if an item is a constant size regardless of the child's distance toward it, then I believe this will interfere with understanding of spatial relationships on the gameboard, because items will look as if they are stuck to the screen (since their size is not changing with distance) rather than being attached to the real world. Another approach considered was for the child to only be able to collect lemons when they are a certain distance of the item. For example, items could show as "not collectable" if the child is too far; they would become "collectable" when the child gets close to them, but not too close. This approach would ensure that all items appeared roughly similar size between children. However, this approach was discarded because pilot studies showed that children were confused about the distance thresholds, since it is unclear at what distance the child should be from the item. Finally, there were some approaches used to ensure all children take similar motion paths to collect lemons. The first approach considered was to control the children's start and end movements – for example the game could only start if the child was looking at the temple in the center of the gameboard from a certain distance and certain orientation, as if looking inside a tunnel; then the child would go to collect one lemon, then they needed to bring the lemon back to the center and ensure that they keep the same distance and orientation when dropping the lemon into the temple at the center of the gameboard. This would ensure that all children were oriented in a similar way when collecting lemons. However, I thought this would be frustrating and not fun enough for children to need to properly orient themselves so frequently. Thus, in order to partially control the motion paths between children, I decided that each level would start as the child touched the temple in the middle of the gameboard, and each set of 4 collected lemons would need to be brought to the temple in the middle of the gameboard. Proper distance and orientation was not required in this approach, thus the child was not constrained as much, but I believe this approach helped to create similar motions between different children. I believe for young children (5-6 year olds) it would be difficult to enforce more constraints on orientations and distances to items; however, for older children it may be possible to create games in which these kinds of motion constraints are built into the fun of the gameplay, thus leading to more precise measurements of user motion and performance.

5.4.6 Verify a Broader Range of AR Designs

Children's reactions to the different experimental conditions are influenced not only by the type of AR interaction technique they are required to use, but also by the game environment in which the interaction technique is situated. For instance, the finger selection in the ARC game was situated within a simple game of lemon collection, compared to the paddle selection interaction in the Puppy Plus game which were situated in a more complex game about mathematics. I have exercised caution when drawing comparisons between interaction techniques across different games in this study, and further research is required to studied these different techniques under the same game environment.

A related limitation is the fact that in the ARC study I have compared interaction techniques only within the context of one experimental game. Children's motivation toward the game environment may have mediated their performance and willingness to engage with the game for prolonged period of time. Additionally, the complexity of the game caused certain kinds of issues to emerge, but it is possible that other games may yield different kinds of issues. For example if a game is played in a physical environment involving more physical objects and the gameplay requires memorizing spatial locations, this will likely yield more problems such as occlusion due to physical objects being in front of other objects; spatial cognition issues related to remembering 3D spatial locations, and related to understanding when virtual objects overlap physical objects on the screen; or issues related to manipulation and bumping into physical objects. Because of these design factors, it is possible that children will react differently to the same interaction if it was presented in a different game or different physical context.

Such different game environments can lead to different kinds of issues, and subsequently a different kind of qualitative coding scheme. Thus, the coding scheme and research findings are limited by the experimental scenario, and further studies must compare children's performance and usability issues on the same interaction across different game environments.

5.4.7 Broaden Generalizability and Population Recruitment

Further research is needed to understand how a wider range of children, with varied socioeconomic status (SES) and technical experience, will perform on handheld AR experiences.

Children in the current studies were recruited mostly from the Emory University subject pool via direct contact, and from Georgia Tech via campus fliers or by word of mouth. My informal observations during the ARC study are that the majority of children came from high SES families, as parents were excited about participating in a study for science research. During the anonymous study survey, children's primary caregivers were asked about their own education level and 45% indicated that they have completed an advanced degree beyond a bachelor's degree; this is much higher than the nationwide average of 12% of the population holding advanced degrees (Ryan and Bauman 2016). High socioeconomic status is associated with a variety of benefits to children's development, such as improved access to material resources, higher access to personal technology, higher academic achievement, better health, higher socioemotional development, and higher exposure to activities that may develop spatial skill (Bradley and Corwyn 2002, Levine, Vasilyeva et al. 2005). Because the current sample was biased towards high SES families, the research results are limited in their generalizability, and further research is needed to replicate this study and compare findings with low-SES populations.

5.4.8 Investigate Speed vs. Accuracy Tradeoffs

Future research can further investigate the relationships between children's speed vs. accuracy. In rapid repeated trial tasks, there is typically a tradeoff between speed and accuracy, whereby some participants may choose to be quick at the expense of inaccuracy. In the ARC study, children were instructed to complete the task as quickly as possible, but they were not extremely pressured to perform too quickly. The lemon collection task was not reset if a child performed an inaccurate selection, because I believed this would add too much frustration to children. However, features were designed into the game to give a sense of urgency: lemons would hide and reappear in a different position on the gameboard if a child took too long to collect them, and a timer would always be visible on the screen. However, the timer did not run out on any children when they played the game, and, after the tutorial level, children typically collected the lemon as soon as it appeared in its initial position on the gameboard.

I analyzed if any correlation existed between children's task completion time and the number of selection errors (see Chapter 4). If children were not paying attention to the task and performing trial-and-error actions, then one would expect a negative correlation, whereby children who take shorter amount of time would have more errors. The analysis showed that a general positive correlation exists, indicating that children who have more errors also took longer to complete the task. The fact that there were no negative correlations between speed and accuracy indicates that, overall, children weren't simply doing trial-and-error tasks. Analyzing this effect within each age group revealed this significant correlation existed only within the 7-8 year-old group. It is unclear why this correlation only existed within the 7-8 age group. Splitting across age groups reduced the power of the correlational analysis, so it may be that a correlation exists but it is not detectable with the reduced number of datapoints; and/or, it may be that within the other age groups there is higher variability, with some children performing inattentively and leaning towards speed-accuracy tradeoffs, while others performing the task attentively. Further research is required to determine under what conditions children exhibit speed-accuracy tradeoffs, and how this impacts child performance measurement studies in augmented reality.

5.4.9 Improve Reliability of the Technology Experience Metric

The metric for measuring technology experience in the ARC study was a question on the parent survey, asking *"Which of the following devices has your child used to play games?"* (possible selections: computers, tablets, cell phones, video game systems, and handheld game systems). This metric has found interesting effects such as the fact that girls used significantly less devices, and negative correlations between technology experience and accuracy. However, there may be validity and reliability issues with this metric. This metric may not be a valid measure of technology experience since higher number of devices may not necessarily mean more expertise, usage, or comfort with the technology. Additionally, this metric may have low reliability because it is asked to the parent, and the parent may not know what the child is exposed to outside of the home, or may feel biased in over- or under-reporting the devices available to their child.

Ideally, multiple metrics should be used to triangulate a child's expertise with technology. One additional survey metric is to ask the child, rather than their parents, about their own usage of devices; in fact this was done in the ARC study through an open-ended interview question, but the young children seemed to have difficulty remembering and/or verbalizing their expertise with devices. Another approach is to ask children/parents about how much time is spent in a typical day using a specific technology, or about child restrictions on use of technology; this was asked also in the ARC study in the parents' questionnaire, but the data was highly variable and no significant patterns were found. A better approach may be to ask this question to the child, since parents may have difficulty estimating the time a child spends using devices, especially when the parent is not with the child, or, they may be biased when reporting this value (Livingstone 2003). Another method, suitable for longitudinal studies, is to use an experience sampling method whereby children can be periodically asked to report on their experience; this metric does not suffer from unreliability due to memory or inability of parents to know what their child is doing, however it is more time consuming for participants (Subrahmanyam, Greenfield et al. 2001). Another approach is to ask children/parents about attitudes towards technology, through rating of statements such as "I can learn a lot of things with a smartphone". This metric can provide an additional metric for triangulation, although research indicates it does suffer from unreliability, for instance girls tend to hold less confidence and less positive perceptions of their own skills with using technology (Cooper and Weaver 2003). Another triangulation method is to gauge a child's expertise by directly measuring their performance on basic tasks within the experimental setting (Volman, van Eck et al. 2005). Future work can use triangulation of multiple such metrics in order to study specifically how previous technology exposure can be beneficial or detrimental to children's use of handheld AR technology.

5.4.10 Improve Reliability of Subjective Experience Instruments

A modified version of the Smileyometer self-report instrument (Read 2008). was used in the ARC study to gauge children's level of fun, enjoyment, and comfort, at each experimental condition. During the study, there was little variation detected in children's self reports. Scores on each metric were similar between age groups thus no age differences could be detected even though children of different ages encountered significantly different problems and performance. Existing research on Smileyometers indicates that this instrument performs best when administered multiple times in order to compare a single child's ratings between different conditions (Read 2008). This was done in the current study in order to compare between selection types and movement types; however, in most cases the instrument did not detect differences between experimental conditions, and children's self ratings remained relatively high between conditions. Additionally, the particular words used in the self report instrument in this study may have low validity. The children's ratings for "Ease of Use" and "Comfort" exhibited very high association for 5-6 and 7-8 year olds (Spearman's rho > 0.800, p<0.001 in both groups), but this association became nonsignificant for 9-10 year olds (Spearman's rho = 0.331, p=0.248). This indicates that the instrument may be measuring different constructs, as children of different ages may be understanding the questions in different ways. Future research should investigate the use of different kinds of subjective measure instruments. For example, researchers have suggested a modified version of the Smileyometer which uses variations only of smiling faces, in order to elicit more nuanced responses from children, or researchers may be able to apply a child-friendly version of the NASA TLX to measure children's subjective experiences in AR.

5.4.11 Replicate with Fewer Experiment-wide Type I errors

Multiple statistical tests were performed throughout the ARC experiment. Type I error per test was set at 0.05 (unless the tests were post-hoc tests, in which Bonferroni correction was used to distribute the Type I error). This signifies a 5% chance that, on tests that detected a significant effect, the effects may not actually exist in the real population. A large number of tests were performed throughout this experiment on the same experimental sample, thus increasing the chance of experiment-wide errors, therefore future research should be conducted to replicate these results in studies with fewer tested variables.

5.4.12 Apply this Research to Novel AR Technologies for Children

The current research has been limited to mostly smartphone-based AR applications which used 2D image tracking. The research results are thus limited to this technological context. When using different kinds of devices or tracking technologies, children will likely encounter different kinds of usability issues and performance. In the following section I will discuss the potential implications of using different AR technologies on young children.

5.4.12.1 Smartphone Camera Placement

A high amount of tracking loss issues were created because children put their fingers in the view of the camera, which was placed on the top left of the back of the device. This kind of issue is less likely to happen in phones that have the camera located in the center of the phone, rather than on the side. With such devices children will be able to hold their hands more comfortably wrapped around the sides of the phone without fear of covering the camera, potentially leading to different kinds of grips and potentially better performance because children will be able to more comfortably control the device. However, children using those kinds of phones may encounter similar tracking loss errors when they hold the phone from the bottom, potentially covering the camera even if it is placed away from the sides.

5.4.12.2 Tablets vs. Smartphones

When AR applications are designed for tablet devices special care should be applied because children will likely not be able to perform finger-based selection. In order to select items with the finger on the screen, children will likely have trouble holding the device with one hand while the other touches the screen. Tablet devices are larger and heavier than smartphones, and this can cause issues especially for young children with small and weaker hands. Because of this, designers may use crosshair interactions. Crosshair appears to be more difficult to use for younger children, thus designers will need to use more care when designing the tutorial levels. Large devices such as touchscreens will also yield higher number of drops, and likely will be more difficult to manipulate thus leading more issues related to aiming the camera away, and to children being unable to play for long periods of time due to endurance.

5.4.12.3 3D Tracking Technologies

Newer handheld devices, such as Google's Tango (Google 2016), use 3D tracking technology to enable augmented reality applications that are anchored to physical objects without the need for a 2D printed image. In such applications, children can use the same interaction techniques as studied in this research, with the difference that the AR tracking works by looking at objects in the natural environment rather than a specific printed paper. Assuming that the AR tracking works flawlessly, I expect that children will perform in similar ways as in the above research. For example, 5-6 year olds are expected to still perform significantly slower than 7-10 year olds and should still have problems orienting themselves in space. However, some issues will disappear. For example, the posture strains observed in older children having to slouch over a table will likely disappear as games are made to be played while maintaining an upright body posture.

Since the tracking technology is different, the current research results related to tracking loss and recovery will likely not be applicable, and further research is needed to understand how children will react to natural feature tracking technology. The technology will still have limitations: for example, some objects may be more easily tracked than others, and some play environments may have boundaries at which the tracking stops working properly. Since a printed image is not required anymore for tracking, one potential issue is that children may not know what is being tracked in the environment, potentially leading to children being unable to understand why tracking may be lost or how to recover it. Another related issue occurs prior to starting a game, as the game must interact with the user in order to ensure that the physical environment is rich enough for 3D tracking, thus the user may need to try different configurations of physical objects to ensure the game can be played robustly. The research question in these instances is: how easily children can understand, create and play within the bounds of a 3D trackable environment?

Another research topic with 3D tracking technologies is related to using real objects in the AR experience. This technology can enable children to potentially use real objects, such as their own toys, to interact with virtual content. This creates the potential for game interactions where a user must manipulate a physical object (e.g., touch the teddy bear with a finger, or move a physical paddle in order to scoop up virtual objects). In such cases, children may be required to hold the AR device with one hand, while interacting with physical objects with the other hand. In my research on the Puppy Plus and Bacteria Snap games, young children had trouble with such interactions that required independent actions with the hands; thus it is likely that young children will have trouble with AR experiences that use real objects.

5.4.12.4 Head-Mounted Displays

Head-Mounted Display (HMD) devices, such as the Meta (Meta 2016) and Microsoft HoloLens (Microsoft 2016), place the AR display on the user's eyes, thus allowing them to interact with an augmented real-world space while using both their hands. Currently virtual reality HMD displays such as the Oculus Rift are not advised for children 12 years old. It is unclear how early elementary-school children react to HMD displays. There are unanswered research questions such as: will children be comfortable wearing such devices? Will children's visual development be adversely impacted by immersion in such environments? When display quality improves, will children be able to distinguish between virtual and real objects once they are immersed in an augmented physical world? Will children be able to understand the spatial relationships between real and physical objects when they are both mixed in an augmented space? What kinds of interactions will be easy to perform by children while wearing HMD displays? Will children encounter safety issues, for example physical issues such as hurting themselves as they move their body expecting to interact with virtual items, or not being aware of the real physical environment that is occluded by virtual items, or psychological issues such as believing that virtual content is actually real?

I expect that children will have an easier time moving around such augmented spaces than compared to using a smartphone, because they will not need to understand the difference between the augmented handheld screen and the non-augmented real world (as is the case when using handheld AR, but not when using HMDs which are directly on the user's eyes). However, it is unclear if children will be able to distinguish between real and virtual objects once the displays improve, and, how they will understand spatial relationships between the physical and virtual spaces. The current research can provide some directions for answering questions related to children's ability to interact with such environments. In an HMD experience, there is a variety of interactions that can be designed into the experience. Children could use a crosshair-style interaction whereby an item is selected after the user rotates their head until the center of their head is pointing at an item; then a hand motion (such as a finger flick) could be used to initiate a selection. Using crosshairs in this HMD scenario and in the handheld AR scenario, is an interaction that combines multiple tasks such as (1) generally aiming the device/head at a specific item, (2) precisely aiming the crosshair at the item, (3) holding the device/head still, and (4) pushing a button or initiating a selection. It is expected that children will have some trouble performing this action, by not being able to precisely aim at items by performing gross and fine motor movements with their head and neck, not being able to hold the head still while aimed at an item of interest, and potentially not remembering how to perform the selection action. More research is required to understand how children will react to this interaction. Another approach to interacting is for children to reach their hand into the augmented space and touch real and virtual items. Such an interaction will be similar to how children interact with real objects in the real world, assuming that children do not have problems with proprioception due to mismatches between their real hand and the system's display of their hand. However, since virtual items may lack tactile feedback and realistic spatial cues, this may yield to children having problems similar to observed in my games of Puppy Plus and Bacteria Snap, whereby children had trouble understanding the spatial relationships between physical and virtual objects, thus leading to improperly moving items in space. Much more research is required to understand how children will react to HMD-enabled AR, and what kinds of interaction techniques are appropriate for children in such environments.

CHAPTER 6

CONCLUSION

Augmented reality (AR) has been shown to have measurable benefits in enriching children's lives, by advancing education through in-situ 3D visualizations, providing entertainment through whole-body interaction, and enhancing physical & cognitive rehabilitation through motivational engagement. Although such experiences were typically confined to desktop computers, the increasing popularity of mobile devices is expected to make AR accessible to large amount of children. In order to realize these benefits, technology designers need to create experiences that are usable by children. Handheld AR interfaces are different from more traditional interfaces, by being small portable windows into physical spaces augmented with digital content, and their use may require users to employ more complex motor and cognitive skills than compared to traditional interfaces. Due to the novelty of handheld AR technology, there are no standard interaction techniques for handheld AR, and little is known about children's ability to use these interfaces.

In the current research, I addressed the following questions: How does children's age relate to performance and usability issues in handheld-AR? How do different handheld-AR interaction techniques compare, in terms of performance and usability issues encountered by children? And, what types of usability issues are experienced by children in handheld-AR?

In order to address these questions, I first constructed several commercial and prototype educational AR games for young children and studied their educational potential, as well as children's ability to use these games. I contributed analyses of how augmented reality can be applied in educational contexts. Further, I generated a usability framework that organizes the usability issues observed in my studies and in existing literature on AR systems for children, discusses relationships between developmental psychology literature and children's AR usability, and provides guidelines for designing AR for children. Finally, I performed a systematic study of children 5-10 years old using handheld augmented reality, as they played a smartphone-based AR game using two AR interaction techniques (finger-based vs. crosshair-based selection) under two movement conditions (tunnels vs. no tunnels). Children's performance and usability problems were analyzed through quantitative and qualitative methods.

This research identified complex relationships between usability metrics and children's age across the elementary-school years. Most metrics showed a general linear improvement with increase in children's age. Age was negatively correlated with overall number of problems encountered by children, number of problems with tracking loss and recovery, and number of problems relating to orienting the body. The 5-6 year old children were sometimes significantly different than the 7-10 year olds as a whole. For example, young children were significantly slower, had more tracking losses while moving, and took longer time to recover tracking. However, the issues encountered by young children were not so severe as to cause the termination of gameplay (although some issues did require significant experimenter engagement to fix), and self-report ratings of fun, ease of use, and comfort, were relatively high (typically beyond 4/5) and were not statistically different between age groups, indicating that young children enjoyed the AR experience even though they encountered more usability issues. Some metrics were similar across all age groups (e.g., children across all age groups had similar number of occurrences of dropping the phone, and needed similar amount of instruction to use the crosshair interaction). Interestingly, other metrics showed an inverse trend of increasing frequency in older children (e.g., older children had more posture strains and more interruptions to scratch themselves).

Children studied in the 5-10 year old range were able to perform the popular interaction techniques of finger selection (selecting by touching virtual items on the screen), or crosshair selection (aiming a crosshair at virtual items and selecting with screen edge buttons). These interactions differed on metrics such as task completion time and ease of use, but they did not differ on other metrics such as accuracy, comfort or fun. Levels involving reorienting the body to aim at occluded items in 3D space caused more issues such as lower accuracy, increased issues with body orientation, and higher tracking loss. Younger children had strong problems with body orientation and tracking loss recovery, thus it is suggested that designers avoid these mechanics. When using other interaction techniques, younger children had strong difficulty manipulating two separate objects with their hands, such as interacting with games requiring them to hold a phone with one hand while moving an AR paddle with the other. Use of different interaction techniques under different movement conditions was linked to different cognitive and physical developmental skills that underlie AR performance (e.g., crosshair-based selection employs skills related to visuomotor precision and spatial relations, while finger-based selection employs skills used in block building activities).

Other factors such as technology exposure, gender, and practice effects were also studied. Higher technology exposure was correlated to more issues such as selection errors, issues with body orientation, and issues with tracking loss due to finger occlusion. These are possibly due to children's mistaken expectations of AR technology. Gender differences were detected, with boys having more technology exposure than girls, and therefore experiencing more issues correlated with technology exposure. Analysis of practice effects showed that during the roughly 30 minutes of gameplay, children improved but only in the speed of selecting virtual items while moving around the game space. This indicates that children could learn how to play the game as well as navigate around the physical space, but they did not show rapid improvements in other metrics such as selection accuracy, dealing with tracking loss and recovery, or manipulating the phone.

This research concludes with a set of guidelines, and discussion of future work. Guidelines for designing handheld AR technology for children aged 5-6, 7-8, and 9-10 years old are presented in Chapter 5.2. Directions for future work are presented in Chapter 5.4. Avenues for future work include: leveraging what is now known to be challenging for children's usability and apply this to the design of challenging and engaging educational games; designing intelligent AR games which measure children's performance and provide the appropriate challenge, potentially developing children's undeveloped skills; researching the long term effects of AR exposure and how children's ability to use this technology may improve with time; further investigating the effects of socioeconomic status, previous non-AR technology experience, and gender; improving research methods for studying AR for children; and researching how to design educational AR applications for young children by using exciting novel technologies such as natural feature tracking and head-mounted AR devices.

APPENDIX A

SUMMARY OF RESEARCH QUESTIONS, HYPOTHESES, METHODS AND RESULTS

Table A.1. Research Questions, Hypotheses, and Methods.

RQ1 : How does children's age relate to performance and usability issues in handheld-AR?

RQ1-1: Does speed of selection differ between age groups?

Hypothesis: Younger children will be slower at performing selection tasks.

Results: Across all movement conditions, 5-6 years old group were significantly slower than the 7-8 and 9-10 year olds; however, no significant differences were found between the 7-8 and 9-10 year olds

Analysis method*: *ANOVA (3-way BWW) on data transformed using reciprocal

**** Normality assumption was not met by original data; the same test was performed on transformed data (using reciprocal, which met normality and homogeneity of variances) as well as on the original data. Significant differences between groups were the same as original ANOVA.***

RQ1-2: Does selection accuracy differ between age groups?

Hypothesis: Younger children will be less accurate at performing selection tasks.

Results: No significant differences in accuracy were found between different age groups.

Analysis method*: *ANOVA (3-way BWW) and nonparametric Kruskal-Wallis H test

**** Normality assumption was not met by original data, so nonparametric tests were performed (Kruskal-Wallis H for between factor). Significant differences between groups were the same as when performing ANOVA (3-way BWW).***

Table A.1. (Continued)

RQ1-3: Does accuracy for AR tracking differ between age groups?

Hypothesis: Younger children will have a higher frequency of AR tracking losses.

Results: When selecting items in conditions requiring whole-body movement (tunnel levels), 5-6 year olds encountered significantly more tracking losses than 7-8 and 9-10 year olds; no significant differences were found between 7-8 and 9-10 year olds. When selecting items in conditions not involving whole-body movement (no-tunnel levels), there were no significant differences found between age groups.

Analysis method*: *ANOVA (3-way BWW) and nonparametric Kruskal-Wallis H test

**** Normality assumption was not met by original data, so nonparametric tests were performed (Kruskal-Wallis H for between factor). Significant differences between groups were the same as when performing ANOVA (3-way BWW).***

RQ1-4: Does speed of AR tracking recovery differ between age groups?

Hypothesis*:* Younger children will be slower at recovering the AR tracking.

Results*:* Overall, 5-6 year old children were significantly slower at recovering tracking than compared to 7-8 and 9-10 year olds; there was no significant difference between 7-8 and 9-10 year olds.

Analysis method*: *ANOVA (3-way BWW) and nonparametric Kruskal-Wallis H test

**** Normality assumption was not met by original data, so nonparametric tests were performed (Kruskal-Wallis H for between factor). Significant differences between groups were the same as when performing ANOVA (3-way BWW).***

Table A.1. (Continued)

RQ1-5: Does children's self-reported fun, ease-of-use, and comfort change with age?

Hypothesis1*:* As age increases, there will be an increase in self reported ease-of-use.

Results*:* Ease-of-use did not significantly differ between age groups.

Analysis method**:** ***Nonparametric Kruskal-Wallis H test***

Hypothesis2*:* As age increases, there will be an increase in self reported comfort.

Results*:* Comfort did not significantly differ between age groups.

Analysis method**:** ***Nonparametric Kruskal-Wallis H test***

Hypothesis3*:* It is unclear if fun would change across age groups.

Results*:* Fun did not significantly differ between age groups.

Analysis method**:** ***Nonparametric Kruskal-Wallis H test***

Table A.1. (Continued)

RQ1-6: Do usability issues differ between age groups? ***Hypothesis***: Younger children will experience higher number of usability issues. ***Results***: There is a significant correlation indicating that the number of usability issues decreases with children's age. This occurs in multiple issues, mostly related to tracking loss and body orientation. Some issues show negative correlation, indicating that number of occurrences increase with children's age. This occurs in issues such as posture strains and scratching. ***Analysis method*: *Nonparametric Spearman correlations, and ANOVA (3-way BWW) / nonparametric Kruskal-Wallis H tests*** **** Normality assumption was not met by original data, so nonparametric tests were performed. If significant differences between groups were the same as when performing ANOVA analysis (3-way BWW), then the ANOVA analysis was also reported.***
RQ2: How do different handheld-AR interaction techniques compare, in terms of performance and usability issues encountered by children?
RQ2-1: Does speed of selection differ between interaction techniques? ***Hypothesis1***: Interaction techniques that involve independent hand movements will lead to lower speed. ***Results***: The opposite result was found - selecting items in the Finger condition (independent hand movement) was significantly faster than in the Crosshair conditions. ***Hypothesis2***: Interaction techniques that involve whole-body movement will lead to lower speed. ***Results***: Selecting items in the Tunnel conditions (whole body movement) was significantly faster than the No Tunnel conditions.

Table A.1. (Continued)

Analysis method: ANOVA 3-way (BWW) and nonparametric Wilcoxon signed-rank test **** Normality assumption was not met by original data, so nonparametric tests were performed. Significant differences between groups were the same as original ANOVA.***
RQ2-2: Does selection accuracy differ between interaction techniques? ***Hypothesis1****:* Interaction techniques that involve independent hand movements will lead to lower accuracy. ***Results****:* No significant difference was found between interactions involving independent movements (Finger selection) vs. coordinated movements (Crosshair selection). ***Hypothesis2****:* Interaction techniques that involve whole-body movement will lead to lower accuracy. ***Results****:* Overall, interaction techniques that involve whole-body movement (Tunnel levels) led to significantly higher errors than non-whole-body movement conditions (No Tunnel levels). ***Analysis method: ANOVA 3-way (BWW) and nonparametric Wilcoxon signed-rank test*** **** Normality assumption was not met by original data, so nonparametric tests were performed. Significant differences between groups were the same as original ANOVA.***

Table A.1. (Continued)

RQ2-3: Does accuracy for AR tracking differ between interaction techniques?

Hypothesis1: Interaction techniques that involve independent hand movements will lead to higher frequency of tracking losses.

Results*:* No significant difference was found between interactions involving independent movements (Finger selection) vs. coordinated movements (Crosshair selection).

Hypothesis2*:* Interaction techniques that involve whole-body movement will lead to higher frequency of tracking losses.

Results*:* Overall, interaction techniques that involve whole-body movement (Tunnel levels) led to significantly higher frequency of tracking losses than non-whole-body movement conditions (No Tunnel levels). For 5-6 and 9-10 year old children, such a difference between conditions was statistically significant; for 7-8 year olds, the difference between conditions was marginally not significant.

Analysis method: ANOVA 3-way (BWW) and nonparametric Wilcoxon signed-rank test

**** Normality assumption was not met by original data, so nonparametric tests were performed. Significant differences between groups were the same as original ANOVA.***

Table A.1. (Continued)

RQ2-4: Does speed of AR tracking recovery differ between interaction techniques? ***Hypothesis1**:* Interaction techniques that involve independent hand movements will lead to lower tracking recovery speed. ***Results**:* No significant difference was found between interactions involving independent movements (Finger selection) vs. coordinated movements (Crosshair selection). ***Hypothesis2**:* Interaction techniques that involve whole-body movement will lead to lower tracking recovery speed. ***Results**:* No significant difference was found between interactions involving whole-body movement (Tunnel levels) vs. no whole-body movement (No Tunnel levels) ***Analysis method: ANOVA 3-way (BWW) and nonparametric Wilcoxon signed-rank test*** **** Normality assumption was not met by original data, so nonparametric tests were performed. Significant differences between groups were the same as original ANOVA.***
RQ2-5: Does child development correlate with performance under different interaction techniques? ***Hypothesis1***: Performance on interaction techniques that involve independent hand movements will be inversely correlated to tests of fine motor skills and physical manipulation.

Table A.1. (Continued)

	Results: When independent hand movement (finger selection) was required, the physical manipulation test (block construction) was inversely correlated with performance on completion time and accuracy on the whole-body movement level. When coordinated hand movement was required (crosshair selection), the fine motor skills test (visuomotor precision) was inversely correlated with performance on selection accuracy and completion time on the non whole-body movement level; and with performance on number of tracking losses on the whole-body movement level. The physical manipulation test (block construction) was also inversely correlated with coordinated hand movement performance on time to recover tracking on the whole-body movement level. The 2D spatial skills test (Spatial Relations) was significantly inversely correlated with coordinated hand movement performance on time to recover tracking on the whole body-movement level. ***Hypothesis2***: Performance on interaction techniques that involve whole-body movement will be inversely correlated to tests of fine motor skill (visuomotor precision test), physical manipulation skill (block construction test), and spatial relations skill (2D spatial relations test). ***Results***: When whole-body movement was required (tunnel conditions), the physical manipulation skill was inversely correlated with performance on independent hand movement conditions (on the metrics of completion time and accuracy); the fine motor skill was inversely correlated with performance on coordinated hand movement conditions (on metrics of number of tracking losses); finally, the spatial relations skill was inversely correlated with performance on coordinated hand movement conditions (on the metric of time to recover tracking). When whole-body movement was not required, the fine motor skill was inversely correlated with performance on coordinated hand movement (on metrics of task completion time and number of errors); in the same condition, the physical manipulation skill was inversely correlated (to metric of time to recover tracking).

Table A.1. (Continued)

Analysis method: Nonparametric Spearman correlations calculated between Developmental Test Scores and AR Performance Metrics, after the effects of Age and Technology Experience have been removed

**** Normality assumption was not met by original data, so nonparametric correlation was performed.***

RQ2-6: Does children's self-reported fun, ease-of-use, and comfort change between interaction techniques?

***Hypothesis1**:* Interaction techniques requiring independent hand movements or whole-body movements will yield less self-reported ease-of-use.

***Results**:* Ease of use was reported significantly higher for the interaction technique requiring independent hand movements (Finger selection).

Analysis method: Nonparametric Wilcoxon signed-rank test

***Hypothesis2**:* Interaction techniques requiring independent hand movements or whole-body movements will yield less self-reported comfort.

***Results**:* Comfort did not significantly differ between interaction technique conditions.

Analysis method: Nonparametric Wilcoxon signed-rank test

Table A.1. (Continued)

***Hypothesis3**:* It is unclear if fun would change across interaction techniques.

***Results**:* Fun did not significantly differ between interaction technique conditions.

Analysis method: Nonparametric Wilcoxon signed-rank test

RQ2-7: Do usability issues differ between interaction techniques?

Hypothesis: Interaction techniques that involve independent hand movements will lead to more usability issues.

Results: No significant differences were found between Finger vs Crosshair selection techniques; however, descriptive statistics show some trends.

Hypothesis: Interaction techniques that involve movement difficulty will lead to more usability issues.

Results: Significant differences were found in several issues related to tracking loss and body movement, whereby more issues occurred in Tunnel than No Tunnel levels.

Analysis method*: *ANOVA (3-way BWW) and nonparametric Wilcoxon signed-rank tests or Sign tests

**** Normality assumption was not met by original data, so nonparametric tests were performed. If nonparametric normality assumption was not met for Wilcoxon test, then Sign test was used. If significant differences between groups were the same as when performing ANOVA analysis (3-way BWW), then the ANOVA analysis was reported.***

Table A.1. (Continued)

RQ3: What types of usability issues are experienced by children in handheld-AR? ***Hypothesis***: Children will experience problems that can be linked to developing areas of physical and cognitive skills ***Results***: The types of problems encountered by children in my research are listed Table 4.4.4.. ***Analysis method: Generated via qualitative video coding, with coding scheme created in this research (reliability: Kappa > 0.8)***

APPENDIX B

VIDEO CODING SCHEME

Behavior Coding Scheme

PASS 1

Category: Codes during Tutorial

- THLP_FINGER – Experimenter needs to give help about how to select with finger
- THLP_TUNNELS~ - Experimenter needs to give help about looking into the tunnels (first time only)
- THLP_CROSSHAIR~ - Experimenter needs to give help about how to use the crosshair (first time only)
- SPACEV~ / SPACENV~ *(as below)*
- LEVELSEC_S / LEVELSEC_E *(as below)*
- SAIDU~ *(as below)*
- LIKE~ *(as below)*
- PHONEDROP_* *(as below)*

Category: Game Events

- LEVELSEC_S - Level Started (record on first green spell appearance in level)
- LEVELSEC_E - Level Ended (record on last green spell disappearance in level)
- TL_* – Tracking Lost (>1s)
- *Modifier is one of: W, A, F, C, U, O~*

Category: Help and Interruptions

- HLPME~ - Child asks for help
- HLPV_SINGLE~ - Experimenter gives quick verbal help
- INTH1_S~ /E – Gameplay interrupted: When experimenter helped for a period, just verbally
- INTH2_S~ /E – Gameplay interrupted: When experimenter helped for a period, and had to take away the phone, or move the paper, or touch the child
- IGN~ - Child ignored instructions
- INTC_S~ /E – Gameplay interrupted: Child does something (verbal or nonverbal) which causes their gameplay to be interrupted (e.g., looking at experimenter)
- SAIDU~ – Child or Experimenter said something that is Unknown (can't be understood)

Category: Mental and Emotional States

- SPACEV~ - Child gives indication of space (verbally)
- SPACENV~ - Child gives indication of space (non-verbal)
- FRUV~ - Indication of frustration or **DISLIKE** (verbally)
- FRUNV~ - Indication of frustration or dislike (non-verbal)
- SUGG~ - Child makes a suggestion (verbally)
- CONF~ - Indication of confusion (verbally)

- BOR~ - Indication of boredom (verbally)
- LIKE~ - Indication of liking (verbally)
- AHA~ - Indication of “aha” moment (verbally)
- TIRV~ - Indication of physical tiredness (verbally)

PASS 2

Category: Grips

Grip modifiers are: CRAB, CURL, STR, CORN, BOTTOM, X

- Grip switches
  - R**GSW**_* – Grip of right hand has switched
  - L**GSW**_* – Grip of left hand has switched
  - **RL**GSW_*r_*l– <right,left> Grip of each hand has changed at the same time.
  - **X**GSW_*r_*l– <right,left> Grip state of each hand at the beginning of the level, or after interruption

Category: Backs

Back modifiers are: STR, BENT

- Back switches
  - BSW_* – Back posture has switched (only if held > 3 seconds)
  - XBSW_*– Back posture at the beginning of the level

Category: General Movement

- STEP - Took steps (code for each 2s)
- SCRATCH - Scratches (code for each 2s)
- TIRNV_* - Tiredness, muscle strain, or stiffness observed non-verbally as:
  - HSTR - Fingers / hand / arm stretched
  - HSHAKE - Hand shaken
  - BSTR - Body stretching
  - BSIT – Body sitting down
- ELT - Elbow or hand is resting on table
- PHONEDROP_* - Phone is dropped or slips
  - *Modifiers: PARTIAL, FULL*
- PHONEDOWN - Puts the phone down on table
- BUMP~ - Bumps or trips body into physical object (or trips over themselves)

Symbols

xxx_S/E – must code the start (xxx_S) and end time (xxx_E)
xxx_SINGLE – code this if event xxx happens for <2 seconds
xxx_S/E/SINGLE – if it's short, use _SINGLE; if it's longer use _S and _E

* must be replaced by a modifier (see the code description for a list of modifiers)
~ must be replaced with a description (e.g., about what the child is saying or doing)

Behavior Code Descriptions

What to Code During the Tutorial Level ?

During the tutorial level, code only these things:

- LEVELSEC_S at the first green spell of the tutorial (i.e.: in the 2D gameplay)
- LEVELSEC_E at the last green spell of the tutorial (i.e.: in the AR gameplay)
- THLP_FINGER – Experimenter needs to give help about how to select with Finger
- THLP_TUNNELS~ - Experimenter needs to give help about looking into the Tunnels
- THLP_CROSSHAIR~ - Experimenter needs to give help about how to use the Crosshair
- SPACEV~ / SPACENV~ - If child indicates knowledge of space *(see description below)*
- SAIDU~ - If you can't understand what is being said *(see description below)*
- LIKE~ - If child indicates liking something *(see description below)*
- PHONEDROP_* - If child drops the phone either PARTIAL or FULL *(see description below)*

Category: Game Events

LEVELSEC_S and LEVELSEC_E

When To Code: In all levels.
We need to record a few reference points in order to be able to correlate the video coded behaviors with gameplay data logs. This is done by coding level section starts and ends.

During each level there are a series of green spells (see figure below showing when the spells occur, and what they look like). These occur BEFORE and AFTER each set of 4 lemons, so there are a total of 8 green spells in each level.

LEVELSEC_S is coded on the first spell in each level, and LEVELSEC_E is coded on the last spell in each level. So, in each level only 2 out of 8 green spells get tagged.
When you code LEVELSEC_S, you should also code the state of the child's posture (using XGSW_* and other X-codes)

Warning: Tag these when you hear the SOUND of the spell being cast (we do this because sometimes the gameplay video is not aligned with the front camera).
Warning: The sound of the spell being cast is the same as the sound of child receiving a star (the child receives a star after the first 2 lemons in each level).

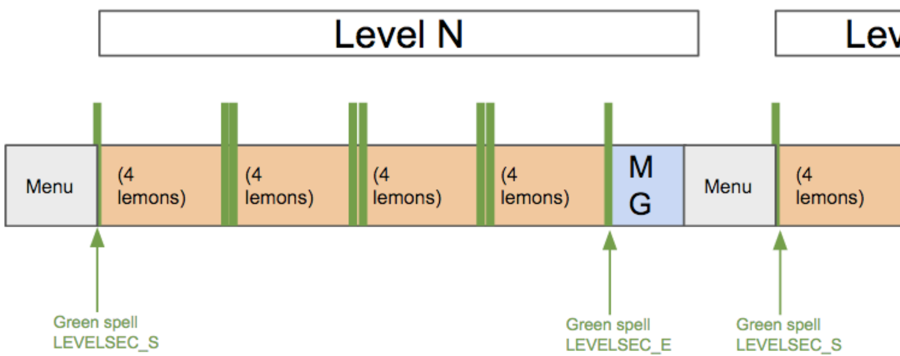


Figure B.1 Structure of one game level.

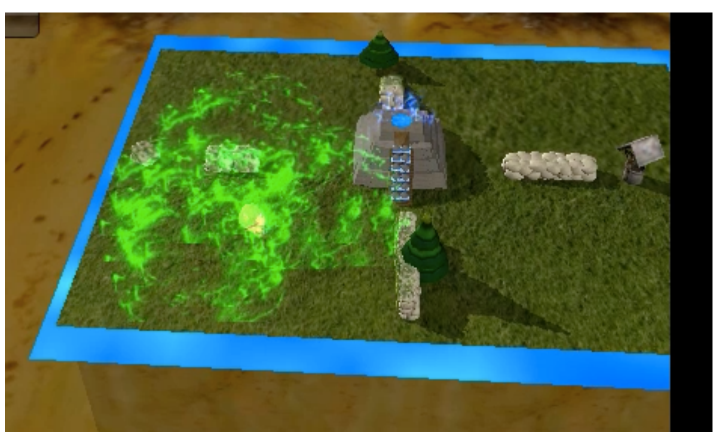


Figure B.2 Green spell at the end of each level section.

TL_* – Tracking Lost

When to code: Only code these things between LEVELSEC_S and LEVELSEC_E of non-tutorial levels.
Record this code when tracking is lost. During gameplay, the camera may stop seeing the paper and cause the tracking to be lost (see figure below). Use this code to record the reason why tracking was lost in the first place. Possible reasons are:

W – Walking: the child was walking and the tracking was lost.
A – Away: while not walking, the child aimed the camera away from the paper.
F – Finger: while not walking, the child put their finger in the way of the camera.
C – Close: while not walking, the child moved the camera too close to the paper.
U – Unknown reason.
O~ – Other reason, often when child is not focusing on game, with explanation, e.g., *"TL_O Dropped the phone", or "TL_O Child is stretching", or "TL_O because frustrated stomping of feet"*

Use the above list to help you disambiguate if tracking happens for multiple reasons. For example, if the child is walking and puts their finger in the way and then gets frustrated and puts the phone down, then just code this as TL_W, because the tracking was lost while the child is walking.

Timing: Only code this if tracking loss lasts more than 1 second (we do this because sometimes the gameplay video shows a quick flash of tracking loss, due to the software issues).

Figure B.3 Gameplay replay when tracking is working (left), and when tracking is lost (right)

Category: Help and Interruptions

When to code: Only code these things between LEVELSEC_S and LEVELSEC_E of non-tutorial levels.
Though experimenters are encouraged to be hands-off in helping participants, the game might present unique challenges that often require extra explanation or help.
HLPME~: Sometimes the child asks for help. When this happens, write what the child

says. Example:

[01:37] HLPME "I'm stuck"

HLPV_SINGLE~: The experimenter might respond with a quick verbal suggestion. Example:

[01:41] HLPV_SINGLE "You have to aim the center of the screen toward the lemon"

INTH1_S~ /E: Interrupted for a period, with help that was only verbal.

INTH2_S~ /E: Interrupted for a period, and had to take away the phone, or move the paper, or touch the child.

Sometimes the experimenter needs to step in, in order to solve the problem. When this happens, the data from the gameplay doesn't show the child's normal gameplay, so we need to keep track of that period. Note: Between INTH_S and INTH_E, don't code anything else.

Relation to other codes: If the child's grip is different after the interruption, then also code XGSW at the time of interruption end.

IGN~: Children might ignore instructions from the experimenter or from the game. Example:

[02:29] IGN child did not follow instructions about holding crosshair

INTC_S~ /E: Child did something non-game related which interrupted the gameplay. The child's gameplay might be interrupted because they do something unrelated to the game. For example they might turn away to talk to their parent, or might be distracted by items on the table, etc. Example:

[04:35] INTC_S child looks around the room

[04:40] INTC_E

Relation to other codes: If the child's grip is different after the interruption, then also code XGSW at the time of interruption end.

SAIDU~: Child or experimenter said something that can't be understood. When annotating videos, you might not understand what is being said. If that happens, you need to flag this, so we can analyze it later.

Category: Mental and Emotional States

When to code: Code these things between LEVELSEC_S and LEVELSEC_E, and also at the beginning of each level.

We can get a lot of information from children's mental and emotional states by observing their behaviors. These codes are designed to detect a variety of verbal and nonverbal behaviors.

SPACEV~: How are children thinking about the space where the game occurs? Do they think of the game world as being on the screen, or as being on top of the paper? We can understand this by observing how the child talks about space. Therefore it is important to record every time the child talks about any of these things:

- the phone viewing angle
    - *[09:30] SPACEV "I'm very high"*
- the location of the paper, the game world, or game objects
    - [09:13] SPACEV "The cat is behind the tunnel"
- the way the gameworld feels real
    - [09:11] SPACEV "Woa I'm cold"
- movement or spatial directions
    - [08:31] SPACEV "I have to back up"
    - [12:33] SPACEV "I don't understand why the tunnels are moving"

SPACENV~: The child might refer to space through other, non-verbal means. For example, the child might rearrange the objects on the table; or, they might look back-and-forth from the phone to the paper; etc. These kinds of behaviors are non-verbal indications that the child is thinking about the space.

FRUV~ Frustration may be noticed verbally or nonverbally. Verbal remarks that may sound like, "I'm not having fun", "This isn't fun", "I don't like this game", "I hate this game", "This is impossible", and "I can't do this".

FRUNV~ Non-verbal observations of frustration are for example: frustrated repeated tapping on the screen, growling/groaning when tracking is lost, etc.

SUGG~: Children can make suggestions of how to improve the game, and this should be recorded. Example:

[11:32] SUGG "I think these lemons should be green"

CONF~: Children might become confused about something. Record this only when there is a verbal indication. Examples: "I am so confused", "What am I supposed to do here?", "I don't get this".

BOR~ - Participants may also become bored and say so out loud. "This is boring" or "I'm bored" type statements that signify a player is becoming actively disengaged and disinterested in play.

LIKE~ - Participants might indicate that they like something. "This game is awesome", "I like these tunnels, they are hard but not too hard", "The cat is so cute", etc.

AHA~ - Players will also experience "Aha!" moments where they suddenly understand how to do something in the game. You may hear "Oooooh I get it now!", "I can't believe I didn't understand that before", "I understand", and "I see!" as well as any other statements that signify a participant suddenly understands something that previously eluded them.

TIRV~ - Players might get tired, and this might be verbal or non-verbal. Use code TIRV if the player says they are tired, e.g., “My back hurts”, or “I’m tired from standing so much”, etc. If the tiredness is non-verbal, use the TIRNV_*_ code (see below).

Category: General Movement

When to code: Only code these things between LEVELSEC_S and LEVELSEC_E of non-tutorial levels.

Body movement needs to be coded because we can learn if children are having trouble (e.g., are they getting tired? being clumsy?); and because body movement observations can shed more light on why the gameplay data has been logged in a certain way (e.g., if the gameplay data indicates shaking, is it because the person is shaking the phone while walking, or shaking the phone while standing still?).

STEP: This code helps us to understand when and for how long children take steps. Use this code whenever a child is taking steps. You don’t need to log each step – just code based on the duration of the child’s movement. If the child is taking steps for 2 seconds or less, then you just need to code one STEP. If the child is moving for longer, then repeat this code as necessary (e.g., if the child moves 4 seconds, then code two STEPs).

SCRATCH: The child scratches themselves. Use this code for every 2 seconds of scratching (e.g., if the child scratches 4 seconds, then code two SCRATCHes).
Relation to other codes:

- If you see that the child stops paying attention to the game while scratching, then also code INTC.
- If the child was holding the phone before scratching: If the child’s grip changed after scratching, use GSW.

ELT: Code this when you see that the child rests their elbow or hand on the table. This could happen because the child is tired, or because the child wants to steady themselves, or etc.

TIRNV_*_S/E/SINGLE – Use this code to record when the child shows signs of tiredness, strain, or stiffness. When someone is having a strain on their hand muscles, they might do non-verbal behaviors like stretching their fingers for a bit, or shaking their hand. If the behavior lasts less than 3 seconds, then use the _SINGLE modifier. If the behavior lasts longer, use _S and _E to denote the period start and end. The other modifier is one of the following:

- HSTR: Child stretches their hand or fingers

- HSHAKE: Child shakes their hand quickly
- BSTR: Child stretches their body (might be just the back or arms)
- BSIT: Child sits down

Relation to other codes:

- If you see that the child stops paying attention to the game while doing this, then also code INTC.
- If the child's grip changed after doing this, use GSW.

PHONEDROP_ - For some children it's hard to hold the phone, and this can cause the child to lose control of the phone, and the phone might slip from their hands. Use PHONEDROP_FULL when you see that the phone has completely slipped away from the child and dropped on the table or floor. Use PHONEDROP_PARTIAL if the phone slipped but the child has caught it before it reached the table or floor.

PHONEDOWN – Another indicator of tiredness or unwillingness to play is when the child puts the phone down.

BUMP~ – Children might not pay attention to their environment when they play the game. This can cause them to bump into objects, or fall off objects. For example, while playing in the tunnel levels, the child might bump their stomach on the table; or, the child must bump their leg on the feet of the table. Whenever the child collides with another physical object, this should be recorded.

[19:32] BUMP child hits table with foot

Category: Grips

When to code: Only code these things between LEVELSEC_S and LEVELSEC_E of non-tutorial levels.

GSW: Grip Switches

While playing the game, children will hold the phone in many different ways. They will start playing the game holding the phone in a certain way, and during gameplay they may switch the grip.

By analyzing the time when grip switches happen, we can determine how long the children use each grip. You must code RGSW_* when you see the grip change on the right hand, and LGSW_* when the grip changes on the left hand.

XGSW_*_*: At the beginning of each level section start (ie: when you code LEVELSEC_S), you must code XGSW_*_* so we know what the first grip is. When To Code: Code this always right before LEVELSEC_S. And, if the grip changed during an interruption, then code this before the _E end of the interruption.

Only code if the grip changes from one grip type into another type.

Category: Backs

When to code: Only code these things between LEVELSEC_S and LEVELSEC_E of non-tutorial levels.

BSW: Back Switches

This code measures the children's middle/lower-back posture. By analyzing the times when children change their back posture, we can determine how long the children stay in unnatural postures such as tilting or leaning. You must code BSW_* when you see the child change their back posture for longer than 3 seconds.

XBSW_*: At the beginning of each level section start (i.e.: when you code LEVELSEC_S), you must code XBSW_* so we know what the first back posture is. When To Code: Code this always right before LEVELSEC_S. And, if the back posture changed during an interruption, then code this before the _E end of the interruption.

Participant Video Data

Video data comes from 3 cameras + one gameplay screen.

Figure B.4 Video recording collage from different experimental cameras.

Gameplay Structure

The game starts with Level 1: Tutorial, followed by 4 levels.

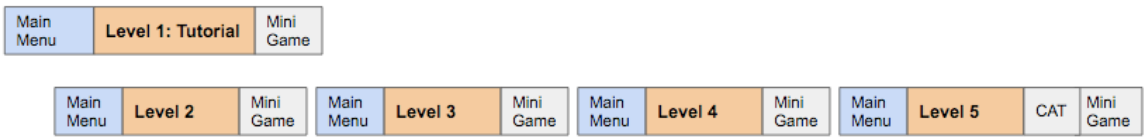


Figure B.5 Structure of one gameplay session.

<u>The Tutorial Level</u> has 3 general sections: It starts with a 2D gameplay, followed by looking around, followed by 3 sections of lemon collection.

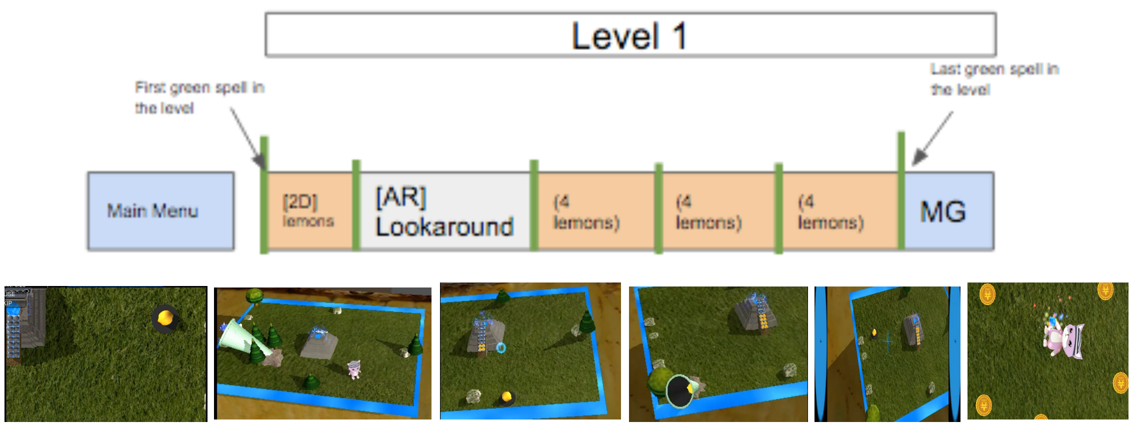


Figure B.6 Structure of the tutorial level.

Regular Levels require the player to collect 4 sets of 4 lemons; they end with a non-AR minigame.

Figure B.7 Structure of a regular game level.

The Last Level is the same as a Regular Level, but before the minigame the player must do a look-around cat task, like in the tutorial.

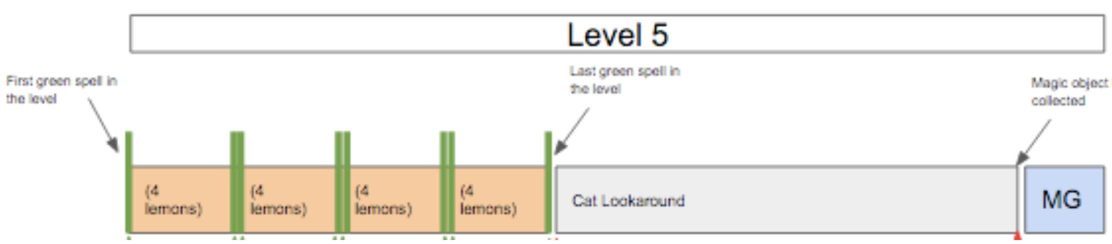


Figure B.8 Structure of the last level of the game.

APPENDIX C

ARC STUDY DESCRIPTIVE STATISTICS

C.1. PERFORMANCE METRICS

Table C.1 Descriptive statistics for Task Completion Time (seconds, per item collected)

Descriptive Statistics For Completion Time (TNT_)				
Age Group		N	Mean	Std. Deviation
AGE 5-6	Finger Selection	14	3.9879	0.8964
	Crosshair Selection	14	5.1853	1.34936
	No Tunnels	14	3.0442	0.70353
	Tunnels	14	6.129	1.56735
	Overall	14	4.5866	1.0158
AGE 7-8	Finger Selection	14	2.6528	0.55248
	Crosshair Selection	14	3.3893	0.41773
	No Tunnels	14	1.8441	0.34303
	Tunnels	14	4.1981	0.60718
	Overall	14	3.0211	0.39771
AGE 9-10	Finger Selection	14	2.4485	0.40171
	Crosshair Selection	14	3.5783	0.65304
	No Tunnels	14	1.9054	0.3765
	Tunnels	14	4.1214	0.72348
	Overall	14	3.0134	0.47324
ALL	Finger Selection	42	3.0298	0.93811
	Crosshair Selection	42	4.051	1.19708
	No Tunnels	42	2.2646	0.74242
	Tunnels	42	4.8162	1.39482
	Overall	42	3.5404	1.00449

Table C.2 Descriptive statistics for Number of Tracking Losses (per item collected)

Descriptive Statistics for Number of Tracking Losses (ERRNTL)				
Age Group		N	Mean	Std. Deviation
AGE 5-6	Finger Selection	14	0.1351	0.08323
	Crosshair Selection	14	0.1349	0.10375
	No Tunnels	14	0.0528	0.05893
	Tunnels	14	0.2173	0.12915
	Overall	14	0.135	0.07904
AGE 7-8	Finger Selection	14	0.0739	0.1047
	Crosshair Selection	14	0.0512	0.08429
	No Tunnels	14	0.032	0.05657
	Tunnels	14	0.0931	0.1251
	Overall	14	0.0625	0.08482
AGE 9-10	Finger Selection	14	0.0452	0.06358
	Crosshair Selection	14	0.0437	0.06119
	No Tunnels	14	0.0166	0.02308
	Tunnels	14	0.0724	0.08092
	Overall	14	0.0445	0.04904
ALL	Finger Selection	42	0.0847	0.09162
	Crosshair Selection	42	0.0766	0.09275
	No Tunnels	42	0.0338	0.0501
	Tunnels	42	0.1276	0.12853
	Overall	42	0.0807	0.08119

Table C.3 Descriptive statistics for Time to Recover Tracking (seconds, per tracking loss)

Descriptive Statistics for Time to Recover Tracking (ERRNTLS)				
Age Group		N	Mean	Std. Deviation
AGE 5-6	Finger Selection	13	2.6588	1.14054
	Crosshair Selection	13	2.3192	1.42247
	No Tunnels	8	2.4595	1.50245
	Tunnels	14	3.2171	1.64618
	Overall	14	2.3112	0.91279
AGE 7-8	Finger Selection	9	1.5842	1.40804
	Crosshair Selection	7	1.184	0.79503
	No Tunnels	7	1.4013	1.67082
	Tunnels	8	1.5921	0.79171
	Overall	10	1.1273	0.87494
AGE 9-10	Finger Selection	7	0.85	0.45257
	Crosshair Selection	6	2.1125	1.94527
	No Tunnels	5	0.514	0.34831
	Tunnels	9	1.7839	1.7054
	Overall	9	1.0347	0.87593
ALL	Finger Selection	29	1.8887	1.32013
	Crosshair Selection	26	1.9659	1.45637
	No Tunnels	20	1.6028	1.54101
	Tunnels	31	2.3816	1.64172
	Overall	33	1.6043	1.06141

Table C.4 Descriptive statistics for Number of Selection Errors (per item collected)

Descriptive Statistics for Number of Selection Errors (ERRA)				
Age Group		N	Mean	Std. Deviation
AGE 5-6	Finger Selection	14	0.9569	0.48781
	Crosshair Selection	14	1.0443	0.83072
	No Tunnels	14	0.7695	0.64042
	Tunnels	14	1.2317	0.59428
	Overall	14	1.0006	0.59599

Table C.4. (Continued)

AGE 7-8	Finger Selection	14	0.8212	0.86475
	Crosshair Selection	14	1.1206	1.13232
	No Tunnels	14	0.5403	0.54384
	Tunnels	14	1.4015	1.32838
	Overall	14	0.9709	0.90006
AGE 9-10	Finger Selection	14	0.7577	0.69531
	Crosshair Selection	14	0.4833	0.37085
	No Tunnels	14	0.4461	0.43525
	Tunnels	14	0.795	0.63325
	Overall	14	0.6205	0.50958
ALL	Finger Selection	42	0.8453	0.68769
	Crosshair Selection	42	0.8828	0.86698
	No Tunnels	42	0.5853	0.55026
	Tunnels	42	1.1427	0.93032
	Overall	42	0.864	0.69451

C.2. SUBJECTIVE MEASURES

Table C.5 Descriptive statistics for Fun (out of 5, per level)

Age Group		N	Mean	Std. Deviation
AGE 5-6	Overall	14	4.5536	0.56482
	Crosshair Selection	14	4.5	0.70711
	Finger Selection	14	4.6071	0.52545
	No Tunnels	14	4.5357	0.57057
	Tunnels	14	4.5714	0.8052
AGE 7-8	Overall	14	4.1964	0.75434
	Crosshair Selection	14	4.1429	0.8419
	Finger Selection	14	4.25	0.72722
	No Tunnels	14	4.25	0.82625
	Tunnels	14	4.1429	0.74495

Table C.5. (Continued)

AGE 9-10	Overall	14	4.0536	0.99604
	Crosshair Selection	14	4.0714	1.0351
	Finger Selection	14	4.0357	1.02777
	No Tunnels	14	4.2143	0.91387
	Tunnels	14	3.8929	1.14654
ALL	Overall	42	4.2679	0.8009
	Crosshair Selection	42	4.2381	0.87121
	Finger Selection	42	4.2976	0.8044
	No Tunnels	42	4.3333	0.77826
	Tunnels	42	4.2024	0.93743

Table C.6 Descriptive statistics for Ease of Use (out of 5, per level)

Age Group		N	Mean	Std. Deviation
AGE 5-6	Overall	14	4.1607	0.75706
	Crosshair Selection	14	4.0714	0.87392
	Finger Selection	14	4.25	0.77831
	No Tunnels	14	4.25	0.80264
	Tunnels	14	4.0714	0.8052
AGE 7-8	Overall	14	4.2143	0.69929
	Crosshair Selection	14	4.0357	0.92952
	Finger Selection	14	4.3929	0.56086
	No Tunnels	14	4.3571	0.66299
	Tunnels	14	4.0714	0.8052
AGE 9-10	Overall	14	4.1964	0.58981
	Crosshair Selection	14	3.8929	0.78883
	Finger Selection	14	4.5	0.51887
	No Tunnels	14	4.25	0.61237
	Tunnels	14	4.1429	0.66299

Table C.6. (Continued)

ALL	Overall	42	4.1905	0.66902
	Crosshair Selection	42	4	0.84824
	Finger Selection	42	4.381	0.62283
	No Tunnels	42	4.2857	0.68202
	Tunnels	42	4.0952	0.74275

Table C.7 Descriptive statistics for Comfort (out of 5, per level)

Age Group		N	Mean	Std. Deviation
AGE 5-6	Overall	14	3.875	0.91856
	Crosshair Selection	14	3.9643	1.11742
	Finger Selection	14	3.7857	0.89258
	No Tunnels	14	4.0357	1.00889
	Tunnels	14	3.7143	1.20439
AGE 7-8	Overall	14	4.0179	0.84617
	Crosshair Selection	14	3.8929	0.90253
	Finger Selection	14	4.1429	0.8419
	No Tunnels	14	4.0357	0.86523
	Tunnels	14	4	0.94054
AGE 9-10	Overall	14	4.3036	0.68766
	Crosshair Selection	14	4.25	0.82625
	Finger Selection	14	4.3571	0.74495
	No Tunnels	14	4.2143	0.8484
	Tunnels	14	4.3929	0.6557
ALL	Overall	42	4.0655	0.82281
	Crosshair Selection	42	4.0357	0.94606
	Finger Selection	42	4.0952	0.84275
	No Tunnels	42	4.0952	0.89196
	Tunnels	42	4.0357	0.97776

C.3. QUALITATIVE PROBLEM METRICS

Table C.8 Descriptive statistics for Number of Usability Problems in each Age and Severity (frequency per child)

Descriptive Statistics Average problems per child				
Age Group		N	Mean	Std. Deviation
AGE 5-6	Total	14	20.6429	15.59392
	Severity 0	14	4.2143	9.90088
	Severity 1	14	14.4286	10.38278
	Severity 2	14	1	1.83973
	Severity 3	14	1	1.35873
AGE 7-8	Total	14	12.0714	11.118
	Severity 0	14	3.3571	3.60784
	Severity 1	14	8.2857	7.87819
	Severity 2	14	0.3571	0.74495
	Severity 3	14	0.0714	0.26726
AGE 9-10	Total	12	9.9167	6.90794
	Severity 0	12	3.8333	3.8573
	Severity 1	12	5.9167	5.38446
	Severity 2	12	0	0
	Severity 3	12	0.1667	0.38925

Table C.9 Observed usability issues, and statistically significant positive correlations (C+), negative correlations (C-), differences between tunnel and no-tunnel conditions, and other group differences (X).

ISSUE CATEGORY	NUM. CHILDREN WHO ENCOUNTERED THIS ISSUE			STATISTICALLY SIGNIFICANT EFFECTS DETECTED (P<0.05)		
	5-6 years old	7-8 years old	9-10 years old	Age	Movement: tunnels (T) vs. no tunnels (NT)	Grips
Overall				C-		
Manipulation						
Losing tracking while walking	14 (100%)	7 (50%)	6 (50%)	X	T > NT	X
Losing tracking by covering the camera with the finger	10 (71%)	11 (79%)	8 (66%)			
Strained grip	0 (0%)	5 (35%)	2 (16%)			
Dropping the phone	2 (14%)	1 (7%)	2 (17%)			
Strained body posture	0 (0%)	4 (29%)	5 (41%)	C+	T > NT	
Space						
Losing tracking by aiming the camera away from the gameboard	8 (57%)	4 (29%)	1 (8%)	C-		
Losing tracking by aiming the camera too close to the gameboard	2 (14%)	2 (14%)	0 (0%)			
Difficulty orienting body in relation to the gameboard	9 (64%)	7 (50%)	0 (0%)	X	T > NT	
Abstract thinking						
Needing initial instruction on how to use crosshair	10 (71%)	11 (79%)	7 (58%)			
Needing in-game instruction on how to use crosshair	2 (14%)	1 (7%)	0 (0%)			
Not understanding the game storyline	1 (7%)	3 (21%)	1 (8%)			
Not understanding general game mechanics	3 (21%)	2 (14%)	0 (0%)			
Difficulties interpreting tracking loss and recovering tracking	7 (50%)	2 (14%)	2 (16%)	C-	T > NT	
Attention						
Bumping or tripping	5 (36%)	5 (36%)	5 (41%)		T > NT	
Interruption due to self-distraction	3 (21%)	1 (7%)	1 (8%)			
Interruption due to scratching	3 (21%)	8 (57%)	8 (67%)	C+		X

Table C.10 Distribution of occurrences of losing tracking while walking.

	5-6 years old	7-8 years old	9-10 years old
Number of children experiencing this issue	14 (100%)	7 (50%)	6 (50%)
Number of occurrences per affected child Min – Max	1-11	1-9	2-8
Number of occurrences per affected child Median	5	2	2.5
Children experiencing this event at Severity 1	14 (100%)	7 (50%)	6 (50%)
Total instances as Severity 1	73	25	20

Table C.11 Distribution occurrences of children losing tracking by covering the camera with the finger.

	5-6 years old	7-8 years old	9-10 years old
Number of children experiencing this issue	10 (71%)	11 (79%)	8 (66%)
Number of occurrences per affected child Min – Max	1-18	1-11	1-5
Number of occurrences per affected child Median	3	2	1
Children experiencing this event at Severity 1	10 (71%)	11 (79%)	8 (66%)
Total instances as Severity 1	47	38	13

Table C.12 Distribution of occurrences of children showing strained grip.

	5-6 years old	7-8 years old	9-10 years old
Number of children experiencing this issue	0 (0%)	5 (35%)	2 (16%)
Number of occurrences per affected child Min – Max	NA	1-3	2-5
Number of occurrences per affected child Median	NA	1	3.5
Children experiencing this event at Severity 1	0 (0%)	5 (35%)	2 (16%)
Total instances as Severity 1	0	7	7

Table C.13 Distribution of occurrences of dropping the phone.

	5-6 years old	7-8 years old	9-10 years old
Number of children experiencing this issue	2 (14%)	1 (7%)	2 (17%)
Number of occurrences per affected child Min – Max	1-1	1-1	1-1
Number of occurrences per affected child: Median	1	1	1
Children experiencing this event at Severity 3	2 (14%)	1 (7%)	2 (17%)
Total instances as Severity 3	2	1	2

Table C.14 Distribution of occurrences of strained body posture.

	5-6 years old	7-8 years old	9-10 years old
Number of children experiencing this issue	0 (0%)	4 (29%)	5 (41%)
Number of occurrences per affected child Min – Max	NA	1-7	1-14
Number of occurrences per affected child: Median	NA	3	2
Children experiencing this event at Severity 0 (occurrence of back bending)	0 (0%)	3 (21%)	4 (29%)
Children experiencing this event at Severity 1 (occurrence of visible strain)	0 (0%)	3 (21%)	3 (21%)
Total instances as Severity 0 (occurrences of back bending)	0	10	8
Total instances as Severity 1 (occurrences of visible strain)	0	4	15

Table C.15 Distribution of occurrences of losing tracking by aiming away from the gameboard.

	5-6 years old	7-8 years old	9-10 years old
Number of children experiencing this issue	8 (57%)	4 (29%)	1 (8%)
Number of occurrences per affected child Min – Max	1-12	1-3	1-1
Number of occurrences per affected child: Median	1.5	1.5	1
Children experiencing this event at Severity 1	8 (57%)	4 (29%)	1 (8%)
Total instances as Severity 1	23	7	1

Table C.16 Distribution of occurrences of losing tracking by aiming too close to the gameboard.

	5-6 years old	7-8 years old	9-10 years old
Number of children experiencing this issue	2 (14%)	2 (14%)	0 (0%)
Number of occurrences per affected child Min – Max	1-3	1-1	NA
Number of occurrences per affected child Median	2	1	NA
Children experiencing this event at Severity 1	2 (14%)	2 (14%)	0 (0%)
Total instances as Severity 1	4	2	0

Table C.17 Distribution of occurrences of difficulty orienting body in relation to gameboard.

	5-6 years old	7-8 years old	9-10 years old
Number of children experiencing this issue	9 (64%)	7 (50%)	0 (0%)
Number of occurrences per affected child Min – Max	1 - 4	1 - 2	NA
Number of occurrences per affected child Median	2	2	NA
Children experiencing this event at Severity 1 (frustration observed)	4 (29%)	7 (50%)	0 (0%)
Children experiencing this event at Severity 2 (verbal help given)	3 (21%)	1 (7%)	0 (0%)
Children experiencing this event at Severity 3 (experimenter had to move the child or phone)	6 (43%)	0 (0%)	0 (0%)
Total instances as Severity 1 (frustration observed)	8	10	0
Total instances as Severity 2 (verbal help given)	5	1	0
Total instances as Severity 3 (experimenter had to move the child or phone)	7	0	0

Table C.18 Distribution of occurrences of needing initial crosshair instruction.

	5-6 years old	7-8 years old	9-10 years old
Children experiencing this event at Severity 2 (verbal help given)	10 (71%)	11 (79%)	7 (58%)
Total instances as Severity 2 (verbal help given)	10	11	7

Table C.19 Distribution of occurrences of needing in-game crosshair instructions.

	5-6 years old	7-8 years old	9-10 years old
Number of children experiencing this issue	2 (14%)	1 (7%)	0 (0%)
Number of occurrences per child Min - Max	1	1	NA
Children experiencing this event at Severity 2 (verbal help given)	2 (14%)	1 (7%)	0 (0%)
Total instances as Severity 2 (verbal help given)	2	1	0

Table C.20 Distribution of occurrences of being confused about the game storyline.

	5-6 years old	7-8 years old	9-10 years old
Number of children experiencing this issue	1 (7%)	3 (21%)	1 (8%)
Number of occurrences per affected child Min – Max	1	1-2	1
Number of occurrences per affected child Median	1	1	1
Children experiencing this event at Severity 0 (confusion observed)	1 (7%)	3 (21%)	1 (8%)
Total instances as Severity 0 (confusion observed)	1	4	1

Table C.21 Distribution of occurrences of being confused about general game mechanics.

	5-6 years old	7-8 years old	9-10 years old
Number of children experiencing this issue	3 (21%)	2 (14%)	0 (0%)
Number of occurrences per affected child Min – Max	1-2	1-2	NA
Number of occurrences per affected child Median	1	1.5	NA
Children experiencing this event at Severity 0 (confusion observed)	3 (21%)	2 (14%)	0 (0%)
Total instances as Severity 0 (confusion observed)	4	3	0

Table C.22 Distribution of occurrences of difficulties interpreting tracking loss and recovery.

	5-6 years old	7-8 years old	9-10 years old
Number of children experiencing this issue	7 (50%)	2 (14%)	2 (16%)
Number of occurrences per affected child Min – Max	1-6	1-5	1-1
Number of occurrences per affected child Median	2	3	1
Children experiencing this event at Severity 0 (confusion observed)	0 (0%)	0 (0%)	1 (7%)
Children experiencing this event at Severity 1 (frustration observed)	2 (14%)	1 (7%)	1 (7%)
Children experiencing this event at Severity 2 (verbal help given)	3 (21%)	2 (14%)	0 (0%)
Children experiencing this event at Severity 3 (experimenter had to move the child or phone)	5 (36%)	0 (0%)	0 (0%)
Total instances as Severity 0 (confusion observed)	0	0	1
Total instances as Severity 1 (frustration observed)	8	4	1
Total instances as Severity 2 (verbal help given)	5	2	0
Total instances as Severity 3 (experimenter had to move the child or phone)	6	0	0

Table C.23 Distribution of occurrences of bumping or tripping.

	5-6 years old	7-8 years old	9-10 years old
Number of children experiencing this issue	5 (36%)	5 (36%)	5 (41%)
Number of occurrences per affected child Min – Max	1-2	1-3	2-3
Number of occurrences per affected child Median	1	1	2
Children experiencing this event at Severity 1	5 (36%)	5 (36%)	5 (41%)
Total instances as Severity 1	7	8	12

Table C.24 Distribution of occurrences of self-distracted interruptions.

	5-6 years old	7-8 years old	9-10 years old
Number of children experiencing this issue	3 (21%)	1 (7%)	1 (8%)
Number of occurrences per affected child Min – Max	1-30	8-8	6-6
Number of occurrences per affected child Median	2	8	6
Children experiencing this event at Severity 0	3 (21%)	1 (7%)	1 (8%)
Total instances as Severity 0	33	8	6

Table C.25 Distribution of occurrences of scratching interruptions.

	5-6 years old	7-8 years old	9-10 years old
Number of children experiencing this issue	3 (21%)	8 (57%)	8 (67%)
Number of occurrences per affected child Min – Max	1-8	1-3	1-12
Number of occurrences per affected child Median	1	2	3
Children experiencing this event at Severity 0	3 (21%)	8 (57%)	8 (67%)
Total instances as Severity 0	10	16	29

C.4. OTHER VARIABLES

Table C.26 Descriptive statistics for Technology Experience per Gender (num devices per child)

Age Group	Gender	N	Mean	Std. Deviation
OVERALL	Female	20	3.4	1.27321
	Male	22	4.2273	0.86914
AGE 5-6	Female	6	2.8333	1.47196
	Male	8	3.875	1.12599
AGE 7-8	Female	7	3.2857	1.1127
	Male	7	4.5714	0.53452
AGE 9-10	Female	7	4	1.1547
	Male	7	4.2857	0.75593

Table C.27 Descriptive statistics for use of grips by age and hand (% of time used per game)

ChildAge	GripHand	GripType	Mean	Std. Deviation
AGE 5-6	LEFT	BOTTOM	0.043632371	0.194936028
		CORN	0.229872953	0.401865729
		CRAB	0.338965197	0.451246978
		CURL	0.195311872	0.378699393
		STR	0.171423962	0.359243773
		NO GRIP	0.020793643	0.133740075
	RIGHT	BOTTOM	0.018884453	0.13360649
		CORN	0.000526892	0.003942896
		CRAB	0.24854218	0.415866291
		CURL	0.678459344	0.447410979
		STR	0.009565398	0.064834752
		NO GRIP	0.044021738	0.184113809
AGE 7-8	LEFT	BOTTOM	0.003253832	0.024349446
		CORN	0.006134908	0.033314157
		CRAB	0.5746502	0.486928267
		CURL	0.144284353	0.3165734
		STR	0.23596242	0.408778189
		NO GRIP	0.035714286	0.187256335
	RIGHT	BOTTOM	0.071428571	0.259870097
		CORN	0.003423936	0.025622389
		CRAB	0.115304532	0.295749433
		CURL	0.695961514	0.436523056
		STR	0.079537755	0.262727117
		NO GRIP	0.034343693	0.165599141
AGE 9-10	LEFT	BOTTOM	0.051259841	0.210668207
		CORN	0.176563896	0.343509818
		CRAB	0.519174391	0.467238171
		CURL	0.089630362	0.270666475
		STR	0.13889166	0.333651191
		NO GRIP	0.003646518	0.019545842
	RIGHT	BOTTOM	0	0
		CORN	0.038263888	0.186212701
		CRAB	0.283910471	0.444355579
		CURL	0.601121136	0.489764538
		STR	0.019363752	0.13415601
		NO GRIP	0.036507421	0.162629022

Table C.28 Descriptive statistics for use of grips by hand, for all ages (% of time used per game)

GripHand	GripType	Mean	Std. Deviation
L	BOTTOM	0.031788123	0.164071129
	CORN	0.13557192	0.31728154
	CRAB	0.475517706	0.477061385
	CURL	0.145747787	0.328249048
	STR	0.184252732	0.370096623
	NO GRIP	0.02087173	0.13636934
R	BOTTOM	0.031609558	0.174505461
	CORN	0.012861956	0.103740642
	CRAB	0.212519491	0.392135546
	CURL	0.661383641	0.455722173
	STR	0.036995229	0.177896624
	NO GRIP	0.038380127	0.170433812

Table C.29 Descriptive statistics for use of grips by age for both hands (% of time used per game)

ChildAge	GripType	Mean	Std. Deviation
AGE 5-6	BOTTOM	0.031258412	0.166818131
	CORN	0.115199922	0.305444998
	CRAB	0.293753688	0.434339204
	CURL	0.436885608	0.478676113
	STR	0.09049468	0.269514567
	NO GRIP	0.03240769	0.160608187
AGE 7-8	BOTTOM	0.037341202	0.18689095
	CORN	0.004779422	0.029615323
	CRAB	0.344977366	0.462651328
	CURL	0.420122933	0.469944846
	STR	0.157750087	0.350957582
	NO GRIP	0.03502899	0.175963069
AGE 9-10	BOTTOM	0.025629921	0.150402026
	CORN	0.107413892	0.283488256
	CRAB	0.401542431	0.46869649
	CURL	0.345375749	0.470118023
	STR	0.079127706	0.259979166
	NO GRIP	0.02007697	0.11639033

APPENDIX D

ARC STUDY DATA COLLECTION INSTRUMENT

SAMPLE

Level 1

Did you have FUN playing that level ?

Absolutely Not	Not Really	So-So	Yes, a bit	Yes, very much !

Notes: ____

Was that level EASY to play ?

Yes, very much!	Yes, a bit	So-So	Yes, a bit	Yes, very much !

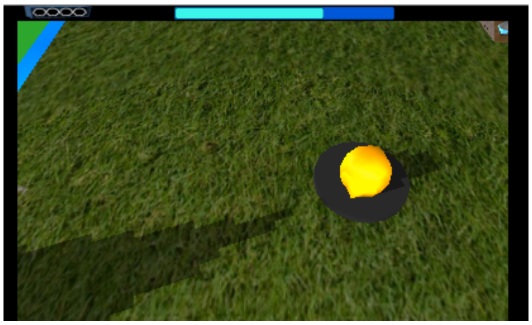
Notes: ____

Were you COMFORTABLE when playing ?

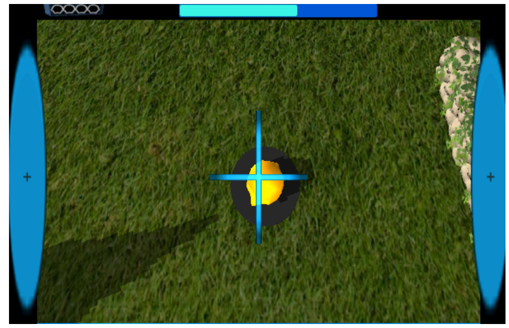
Absolutely Not	Not Really	So-So	Yes, a bit	Yes, very much !

Notes: ____

END OF GAME QUESTIONNAIRE



FINGER



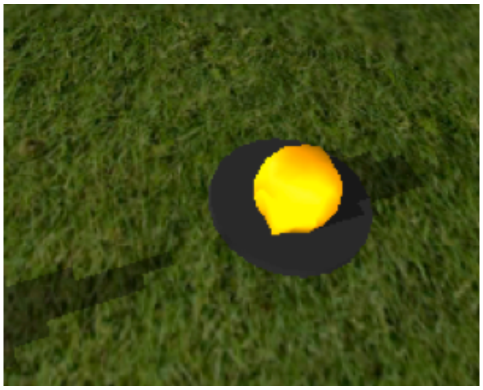
CENTER OF SCREEN

WHICH ONE WAS MORE FUN ?

<table>
<tr><td rowspan="2">FINGER</td><td>BOTH FUN</td><td rowspan="2">CENTER SCREEN</td></tr>
<tr><td>BOTH NOT FUN</td></tr>
</table>

WHICH ONE WAS HARDER ?

<table>
<tr><td rowspan="2">FINGER</td><td>BOTH EASY</td><td rowspan="2">CENTER SCREEN</td></tr>
<tr><td>BOTH HARD</td></tr>
</table>



NO TUNNEL

TUNNEL

WHICH ONE WAS MORE FUN ?

<table>
<tr><td rowspan="2">NO TUNNEL</td><td>BOTH FUN</td><td rowspan="2">TUNNEL</td></tr>
<tr><td>BOTH NOT FUN</td></tr>
</table>

WHICH ONE WAS HARDER ?

<table>
<tr><td rowspan="2">NO TUNNEL</td><td>BOTH HARD</td><td rowspan="2">TUNNEL</td></tr>
<tr><td>BOTH EASY</td></tr>
</table>

Have you ever played a game like this where you play with the phone camera ?

Did you like having to moving around the table ?

Absolutely Not | Not Really | So-So | Yes, a bit | Yes, very much !

Which was your favorite part of this game ?

Which was your least favorite part ?

Do you think your best friend would like this game ?

Absolutely Not | Not Really | So-So | Yes, a bit | Yes, very much !

How old is your best friend's age:

Do you think this game is good for
[ ] younger kids or [ ] older kids or [ ] all ages ?

Should we change anything about this game to make it better ?

REFERENCES

Andersen, T. L., S. Kristensen, B. W. Nielsen and K. Grønbæk (2004). Designing an augmented reality board game with children: the battleboard 3D experience. Proceedings of the 2004 conference on Interaction design and children: building a community, ACM.

Anthony, L., Q. Brown, J. Nias, B. Tate and S. Mohan (2012). Interaction and recognition challenges in interpreting children's touch and gesture input on mobile devices. Proceedings of the 2012 ACM international conference on Interactive tabletops and surfaces, ACM.

Barendregt, W. and M. Bekker (2006). "Developing a coding scheme for detecting usability and fun problems in computer games for young children." Behavior research methods **38**(3): 382-389.

Baumgarten, M. (2003). "Kids and the internet: a developmental summary." Computers in Entertainment (CIE) **1**: 2.

Bekker, M., W. Barendregt, S. Crombeen and M. Biesheuvel (2005). Evaluating usability and challenge during initial and extended use of children's computer games. People and Computers XVIII—Design for Life, Springer**:** 331-345.

Bekker, T. and A. N. Antle (2011). "Developmentally situated design (DSD): making theoretical knowledge accessible to designers of children's technology." Proceedings of the 2011 annual conference on Human factors in computing systems: 2531--2540.

Belmont, L. and H. G. Birch (1963). "Lateral dominance and right-left awareness in normal children." Child Development: 257-270.

Berkovitz, J. (1994). Graphical interfaces for young children in a software-based math curriculum. Conference companion on Human factors in computing systems, ACM.

Billinghurst, M. (2002). "Augmented reality in education." New Horizons for Learning **12**.

Boring, S., D. Ledo, X. A. Chen, N. Marquardt, A. Tang and S. Greenberg (2012). The fat thumb: using the thumb's contact size for single-handed mobile interaction. Proceedings of the 14th international conference on Human-computer interaction with mobile devices and services, ACM.

Bowman, D. A. (2002). "Principles for the Design of Performance-oriented Interaction Techniques." Handbook of Virtual Environments: Design, Implementation, and Applications: 277.

Bowman, D. A. and L. F. Hodges (1999). "Formalizing the design, evaluation, and application of interaction techniques for immersive virtual environments." Journal of Visual Languages and Computing **10**(1): 37-53.

Bowman, D. A., D. B. Johnson and L. F. Hodges (1999). Testbed evaluation of virtual environment interaction techniques. Proceedings of the ACM symposium on Virtual reality software and technology, ACM.

Bowman, D. A., E. Kruijff, J. J. LaViola and I. Poupyrev (2005). 3D user interfaces: theory and practice, Addison-Wesley Reading, MA.

Bradley, R. H. and R. F. Corwyn (2002). "Socioeconomic status and child development." Annual review of psychology **53**(1): 371-399.

Bruckman, A. and A. Bandlow (2002). Human-computer interaction for kids. The human-computer interaction handbook, L. Erlbaum Associates Inc.

Bujak, K. R., I. Radu, B. MacIntyre, R. Catrambone, R. Zheng and G. Golubski (2013). "A Psychological Perspective on Augmented Reality in the Mathematics Classroom." Computers & Education.

Bullens, J., M. Nardini, C. F. Doeller, O. Braddick, A. Postma and N. Burgess (2010). "The role of landmarks and boundaries in the development of spatial memory." Developmental Science **13**(1): 170-180.

Campos, P. and S. Pessanha (2011). "Designing Augmented Reality Tangible Interfaces for Kindergarten Children." Virtual and Mixed Reality-New Trends: 12-19.

Card, S. K., J. D. Mackinlay and G. G. Robertson (1990). The design space of input devices. Proceedings of the SIGCHI conference on Human factors in computing systems: Empowering people, ACM.

Chien-Yu, L., J.-T. Chao and H.-S. Wei (2010). Augmented reality-based assistive technology for handicapped children. Computer Communication Control and Automation (3CA), 2010 International Symposium on, IEEE.

Common Core Standards Initiative. (2012). "Common Core Standards Initiative." Retrieved August 1, 2013, from http://www.corestandards.org/.

Cooper, J. and K. D. Weaver (2003). Gender and computers: Understanding the digital divide, Psychology Press.

Correa, A. G. D., G. A. de Assis, M. Nascimento, I. Ficheman and R. D. Lopes (2007). Genvirtual: An augmented reality musical game for cognitive and motor rehabilitation. Virtual Rehabilitation, 2007, IEEE.

Crucian, G. P. and S. A. Berenbaum (1998). "Sex differences in right hemisphere tasks." Brain and cognition **36**(3): 377-389.

Csikszentmihalyi, M. (1997). Finding flow: The psychology of engagement with everyday life, Basic Books.

De Jong, J. (1957). "The effects of increasing skill on cycle time and its consequences for time standards." Ergonomics **1**(1): 51-60.

De Lisi, R. and J. L. Wolford (2002). "Improving children's mental rotation accuracy with computer game playing." The Journal of genetic psychology **163**: 272-282.

Dempster, F. N. (1981). "Memory span: Sources of individual and developmental differences." Psychological Bulletin **89**(1): 63.

Dunleavy, M., C. Dede and R. Mitchell (2009). "Affordances and Limitations of Immersive Participatory Augmented Reality Simulations for Teaching and Learning." Journal of Science Education and Technology **18**(1): 7-22.

Elliott, D., S. Hansen, J. Mendoza and L. Tremblay (2004). "Learning to optimize speed, accuracy, and energy expenditure: a framework for understanding speed-accuracy relations in goal-directed aiming." Journal of Motor Behavior **36**(3): 339-351.

Fischer, K. and M. Immordino-Yang (2002). "Cognitive development and education: From dynamic general structure to specific learning and teaching." Traditions of scholarship in education. Chicago: Spencer Foundation.

Flavell, J. H., H. Beilin and P. Pufall (1992). "Perspectives on perspective-taking." Piaget's theory: Prospects and possibilities: 107-139.

Freitas, R. and P. Campos (2008). SMART: a SysteM of Augmented Reality for Teaching 2nd grade students. Proceedings of the 22nd British HCI Group Annual Conference on People and Computers: Culture, Creativity, Interaction - Volume 2, Swinton, UK, UK, British Computer Society.

Gallahue, D. L. and J. C. Ozmun (1998). Understanding motor development: Infants, children, adolescents, adults, McGraw-Hill New York.

Geary, D. C. and M. C. DeSoto (2011). "Sex Differences in Spatial Abilities Among Adults from the United States and China." Evolution and Cognition **7**(2): 172-177.

Gelderblom, H. (2009). Designing technology for young children: guidelines grounded in a literature investigation on child development and children's technology. Doctor of Philosophy, University of South Africa.

Good, J., P. Romero, B. du Boulay, H. Reid, K. Howland and J. Robertson (2008). An embodied interface for teaching computational thinking. Proceedings of the 13th international conference on Intelligent user interfaces, ACM.

Google. (2016). "Google Project Tango." Retrieved August 1, 2016, from https://get.google.com/tango/.

Grossman, T., K. Hinckley, P. Baudisch, M. Agrawala and R. Balakrishnan (2006). Hover widgets: using the tracking state to extend the capabilities of pen-operated devices. Proceedings of the SIGCHI conference on Human Factors in computing systems, ACM.

Guven, S., S. Feiner and O. Oda (2006). Mobile augmented reality interaction techniques for authoring situated media on-site. Proceedings of the 5th IEEE and ACM International Symposium on Mixed and Augmented Reality, IEEE Computer Society.

Ha, T. and W. Woo (2010). An empirical evaluation of virtual hand techniques for 3D object manipulation in a tangible augmented reality environment. 3D User Interfaces (3DUI), 2010 IEEE Symposium on, IEEE.

Hancock, M. S. and K. S. Booth (2004). Improving menu placement strategies for pen input. Proceedings of Graphics Interface 2004, Canadian Human-Computer Communications Society.

Hohendorff, B., C. Weidermann, K. Burkhart, P. Rommens, K. Prommersberger and M. Konerding (2010). "Lengths, girths, and diameters of children's fingers from 3 to 10 years of age." Annals of Anatomy-Anatomischer Anzeiger **192**(3): 156-161.

Hornecker, E. and A. Dünser (2007). Supporting Early Literacy with Augmented Books–Experiences with an Exploratory Study. Proceedings of the German Society of Informatics Annual conference (GI-Jahrestagung) 2007.

Hornecker, E. and A. Dünser (2009). "Of pages and paddles: Children's expectations and mistaken interactions with physical-digital tools." Interacting with Computers **21**(1-2): 95-107.

Hourcade, J. P. (2006). Learning from preschool children's pointing sub-movements. Interaction Design and Children.

Hourcade, J. P. (2008). "Interaction design and children." Foundations and Trends in Human-Computer Interaction **1**: 277-392.

Hourcade, J. P., K. B. Perry and A. Sharma (2008). PointAssist: helping four year olds point with ease. Interaction Design and Children.

Hürst, W. and C. Van Wezel (2013). "Gesture-based interaction via finger tracking for mobile augmented reality." Multimedia Tools and Applications **62**(1): 233-258.

Inkpen, K. (1997). Three important research agendas for educational multimedia: Learning, children, and gender. AACE World Conference on Educational Multimedia and Hypermedia, Citeseer.

Inkpen, K. (2001). "Drag-and-Drop versus Point-and-Click Mouse Interaction Styles for Children." Drag-Drop versus Point Click methods.

Inkpen, K., D. Dearman, R. Argue, M. Comeau, C. Ä. Fu, S. Kolli, J. Moses, N. Pilon and J. Wallace (2006). "Left,Äêhanded scrolling for pen,Äêbased devices." International Journal of Human,ÄêComputer Interaction **21**(1): 91-108.

Jackson, L. A., Y. Zhao, A. Kolenic III, H. E. Fitzgerald, R. Harold and A. Von Eye (2008). "Race, gender, and information technology use: the new digital divide." CyberPsychology & Behavior **11**(4): 437-442.

Joiner, R., D. Messer, P. Light and K. Littleton (1998). "It is best to point for young children: A comparison of children's pointing and dragging." Computers in Human Behavior **14**(3): 513-529.

Juan, C., F. Beatrice and J. Cano (2008). An Augmented Reality System for Learning the Interior of the Human Body. International Conference on Advanced Learning Technologies, Santander, Cantabria, Spain.

Juan, C. M., E. Llop, F. Abad and J. Lluch (2010). "Learning Words Using Augmented Reality." 2010 10th IEEE International Conference on Advanced Learning Technologies: 422-426.

Juan, C. M., G. Toffetti, F. Abad and J. Cano (2010). "Tangible Cubes Used as the User Interface in an Augmented Reality Game for Edutainment." 2010 10th IEEE International Conference on Advanced Learning Technologies: 599-603.

Kabbash, P., I. S. MacKenzie and W. Buxton (1993). Human performance using computer input devices in the preferred and non-preferred hands. Proceedings of the INTERACT'93 and CHI'93 conference on Human factors in computing systems, ACM.

Kemp, S. L., M. Korkman and U. Kirk (2001). Essentials of NEPSY assessment, John Wiley & Sons.

Kerawalla, L., R. Luckin, S. Seljeflot and A. Woolard (2006). "“Making it real”: exploring the potential of augmented reality for teaching primary school science." Virtual Reality **10**(3): 163-174.

Kern, D., M. Stringer, G. Fitzpatrick and A. Schmidt (2006). "Curball--A Prototype Tangible Game for Inter-Generational Play."

Kerr, R. (1975). "Movement control and maturation in elementary-grade children." Perceptual and Motor Skills **41**(1): 151-154.

Kesselring, T. and U. Müller (2011). "The concept of egocentrism in the context of Piaget’s theory." New Ideas in Psychology **29**(3): 327-345.

Kim, M. and J. Y. Lee (2016). "Touch and hand gesture-based interactions for directly manipulating 3D virtual objects in mobile augmented reality." Multimedia Tools and Applications: 1-22.

Kruijff, E., J. Swan and S. Feiner (2010). Perceptual issues in augmented reality revisited. Mixed and Augmented Reality (ISMAR), 2010 9th IEEE International Symposium on, IEEE.

Lakatos, D., M. Blackshaw, A. Olwal, Z. Barryte, K. Perlin and H. Ishii (2014). T (ether): spatially-aware handhelds, gestures and proprioception for multi-user 3D modeling and animation. Proceedings of the 2nd ACM symposium on Spatial user interaction, ACM.

Levine, S. C., A. Foley, S. Lourenco, S. Ehrlich and K. Ratliff (2016). "Sex differences in spatial cognition: advancing the conversation." Wiley Interdisciplinary Reviews: Cognitive Science **7**(2): 127-155.

Levine, S. C., J. Huttenlocher, A. Taylor and A. Langrock (1999). "Early sex differences in spatial skill." Developmental psychology **35**(4): 940.

Levine, S. C., M. Vasilyeva, S. F. Lourenco, N. S. Newcombe and J. Huttenlocher (2005). "Socioeconomic status modifies the sex difference in spatial skill." Psychological science **16**(11): 841-845.

Livingstone, S. (2003). "Children's use of the internet: Reflections on the emerging research agenda." New media & society **5**(2): 147-166.

Majid, A., M. Bowerman, S. Kita, D. Haun and S. C. Levinson (2004). "Can language restructure cognition? The case for space." Trends in cognitive sciences **8**(3): 108-114.

Malone, T. W. and M. R. Lepper (1987). "Making learning fun: A taxonomy of intrinsic motivations for learning." Aptitude, learning, and instruction **3**(1987): 223-253.

Marco, J., S. Baldassarri and E. Cerezo (2010). Bridging the gap between children and tabletop designers. Interaction Design and Children.

McKnight, L. and D. Fitton (2010). Touch-screen technology for children: giving the right instructions and getting the right responses. Proceedings of the 9th International Conference on Interaction Design and Children, ACM.

Merians, A. S., D. Jack, R. Boian, M. Tremaine, G. C. Burdea, S. V. Adamovich, M. Recce and H. Poizner (2002). "Virtual reality–augmented rehabilitation for patients following stroke." Physical therapy **82**: 898.

Meta. (2016). "Meta." Retrieved August 1, 2016, from https://www.metavision.com/.

Microsoft. (2016). "Microsoft HoloLens." Retrieved August 1, 2016, from https://www.microsoft.com/microsoft-hololens/en-us.

Morrison, A., A. Oulasvirta, P. Peltonen, S. Lemmela, G. Jacucci, G. Reitmayr, J. N√§s√§nen and A. Juustila (2009). Like bees around the hive: a comparative study of a mobile augmented reality map, ACM.

Mossel, A., B. Venditti and H. Kaufmann (2013). 3DTouch and HOMER-S: intuitive manipulation techniques for one-handed handheld augmented reality. Proceedings of the Virtual Reality International Conference: Laval Virtual, ACM.

Mossel, A., B. Venditti and H. Kaufmann (2013). DrillSample: precise selection in dense handheld augmented reality environments. Proceedings of the Virtual Reality International Conference: Laval Virtual, ACM.

Mott, J., S. Bucolo, L. Cuttle, J. Mill, M. Hilder, K. Miller and R. M. Kimble (2008). "The efficacy of an augmented virtual reality system to alleviate pain in children undergoing burns dressing changes: a randomised controlled trial." Burns : journal of the International Society for Burn Injuries **34**: 803-808.

Newell, A. and P. S. Rosenbloom (1981). "Mechanisms of skill acquisition and the law of practice." Cognitive skills and their acquisition: 1-55.

Nischelwitzer, A., F.-J. Lenz, G. Searle and A. Holzinger (2007). "Some Aspects of the Development of Low-Cost Augmented Reality Learning Environments as Examples." Universal Access in Human-Computer Interaction.: 728-737.

Oda, O. and S. Feiner (2012). 3D referencing techniques for physical objects in shared augmented reality. Mixed and Augmented Reality (ISMAR), 2012 IEEE International Symposium on, IEEE.

Oehl, M., C. Sutter and M. Ziefle (2007). "Considerations on efficient touch interfaces,Äìhow display size influences the performance in an applied pointing task." Human interface and the management of information. Methods, techniques and tools in information design: 136-143.

Olwal, A. and S. Feiner (2003). The flexible pointer: An interaction technique for selection in augmented and virtual reality. Conference supplement of ACM symposium on user interface software and technology.

Payne, V. G. and L. D. Isaacs (2002). Human motor development: a lifespan approach, McGraw-Hill.

Perry, K. B. and J. P. Hourcade (2008). Evaluating one handed thumb tapping on mobile touchscreen devices. Proceedings of graphics interface 2008, Canadian Information Processing Society.

Piaget, J. (1970). "Science of education and the psychology of the child. Trans. D. Coltman."

Pierce, J. S., A. S. Forsberg, M. J. Conway, S. Hong, R. C. Zeleznik and M. R. Mine (1997). Image plane interaction techniques in 3D immersive environments. Proceedings of the 1997 symposium on Interactive 3D graphics, ACM.

Poupyrev, I., M. Billinghurst, S. Weghorst and T. Ichikawa (1996). The go-go interaction technique: non-linear mapping for direct manipulation in VR. Proceedings of the 9th annual ACM symposium on User interface software and technology, ACM.

PTC Inc. (2016). "Vuforia - Home." Retrieved August 1, 2016, from http://www.vuforia.com/.

Public Broadcasting Service. (2011). "PBS KIDS Launches Its First Augmented Reality Mobile App." Retrieved August 3, 2013, from http://www.pbs.org/about/news/archive/2011/fetch-lunch-rush-app/.

Quera, V. and R. Bakeman (2000). GSEQ for Windows: New software for the sequential analysis of behavioral data, with an interface to The Observer. 3rd International Conference on Methods and Techniques in Behavioral Research. Nijmegen, Netherlands.

Quinn, P. C. and L. S. Liben (2008). "A sex difference in mental rotation in young infants." Psychological Science **19**(11): 1067-1070.

Radu, I. (2012). Why should my students use AR? A comparative review of the educational impacts of augmented-reality. Mixed and Augmented Reality (ISMAR), 2012 IEEE International Symposium on, IEEE.

Radu, I. (2014). "Augmented reality in education: a meta-review and cross-media analysis." Personal and Ubiquitous Computing **18**(6): 1533-1543.

Radu, I. and K. Bujak. (2012). "Argon Games: Mountain Rescue and Bacteria Snap." Retrieved August 3, 2013, from http://argon.gatech.edu/games/.

Radu, I., E. Doherty, K. DiQuollo, B. McCarthy and M. Tiu (2015). Cyberchase shape quest: pushing geometry education boundaries with augmented reality. Proceedings of the 14th International Conference on Interaction Design and Children, ACM.

Radu, I., E. Hanlon, Y. Xu, B. Gee and W. Whittaker. (2011). "Puppy Plus." Retrieved August 3, 2013, from http://www.youtube.com/watch?v=ArUp1gxUrOU.

Radu, I. and M. Hewner. (2011). "Spintopia." Retrieved August 3, 2013, from http://www.youtube.com/watch?v=0poxJKNVLIM.

Radu, I. and B. MacIntyre (2009). Augmented-reality scratch: a children's authoring environment for augmented-reality experiences. Proceedings of the 8th International Conference on Interaction Design and Children, ACM.

Radu, I. and B. MacIntyre (2012). Using children's developmental psychology to guide augmented-reality design and usability. Mixed and Augmented Reality (ISMAR), 2012 IEEE International Symposium on, IEEE.

Radu, I., B. McCarthy and Y. Kao (2016). Discovering educational augmented reality math applications by prototyping with elementary-school teachers. 2016 IEEE Virtual Reality (VR), IEEE.

Radu, I., Y. Xu and B. MacIntyre (2013). Embodied metaphor elicitation through augmented-reality game design. Proceedings of the 12th International Conference on Interaction Design and Children, ACM.

Radu, I., R. Zheng, G. Golubski and M. Guzdial (2010). Augmented Reality in the Future of Education. SIGCHI Conference on Human Factors in Computing Systems, Workshop on Educational Technologies and Pedagogy.

Read, J. C. (2008). "Validating the Fun Toolkit: an instrument for measuring children's opinions of technology." Cognition, Technology & Work **10**(2): 119-128.

Revelle, G. and E. Reardon (2009). Designing and testing mobile interfaces for children. Proceedings of the 8th International Conference on Interaction Design and Children, ACM.

Revelle, G. L. and E. F. Strommen (1990). "The effects of practice and input device used on young children‚Äôs computer control." Journal of Computing in Childhood Education **2**(1): 33-41.

Richard, E., V. Billaudeau, P. Richard and G. Gaudin (2007). "Augmented Reality for Rehabilitation of Cognitive Disabled Children: A Preliminary Study." Virtual Rehabilitation: 102-108.

Romeo, G., S. Edwards, S. McNamara, I. Walker and C. Ziguras (2003). "Touching the screen: Issues related to the use of touchscreen technology in early childhood education." British Journal of Educational Technology **34**(3): 329-339.

Rosser, R. A. (1994). Cognitive development: Psychological and biological perspectives, Allyn and Bacon Boston.

Ryan, C. L. and K. Bauman (2016). Educational Attainment in the United States: 2015.

Scaife, M. and R. Bond (1991). "Developmental changes in childrens‚Äô use of computer input devices." Religious Education **69**(1): 19-38.

Scaife, T. M. and A. F. Heckler (2010). "Student understanding of the direction of the magnetic force on a charged particle." American Journal of Physics **78**: 869.

Schell, J. (2014). The Art of Game Design: A book of lenses, CRC Press.

Seo, D. W. and J. Y. Lee (2013). "Direct hand touchable interactions in augmented reality environments for natural and intuitive user experiences." Expert Systems with Applications **40**(9): 3784-3793.

Sesame Workshop. (2013). "Sesame Worskhop and Qualcomm Announce Mobile Technology Collaboration." Retrieved August 3, 2013, from http://www.sesameworkshop.org/our-blog/2013/01/09/sesame-worskhop-and-qualcomm-announce-mobile-technology-collaboration/.

Shelton, B. and N. Hedley (2003). "Exploring a cognitive basis for learning spatial relationships with augmented reality." Technology, Instruction, Cognition and Learning **1**(4): 323-357.

Song, H., J. Clawson and I. Radu (2011). "Updating Fitts' Law to Account for Small Targets." International Journal of Human-Computer Interaction.

Sony Computer Entertainment. "EyePet TM." Retrieved May 13, 2011, from http://www.eyepet.com/.

Strommen, E. (1994). Children's use of mouse-based interfaces to control virtual travel. Proceedings of the SIGCHI conference on Human factors in computing systems: celebrating interdependence, ACM.

Subrahmanyam, K., P. Greenfield, R. Kraut and E. Gross (2001). "The impact of computer use on children's and adolescents' development." Journal of Applied Developmental Psychology **22**(1): 7-30.

Sutter, C. (2007). "Sensumotor transformation of input devices and the impact on practice and task difficulty." Ergonomics **50**(12): 1999-2016.

Swink, S. (2009). "Game feel." A Game Designer's Guide to Virtual Sensation. Burlington, MA: 1.

Tang, A., C. Owen, F. Biocca and W. Mou (2003). "Comparative effectiveness of augmented reality in object assembly." Proceedings of the conference on Human factors in computing systems - CHI '03: 73.

Tanikawa, T., H. Uzuka, T. Narumi and M. Hirose (2015). Integrated view-input interaction method for mobile AR. 3D User Interfaces (3DUI), 2015 IEEE Symposium on, IEEE.

Theng, Y.-L., C. L. Mei-Ling, W. Liu and A. Cheok (2007). Mixed Reality Systems for Learning: A Pilot Study Understanding User Perceptions and Acceptance. International Conference on Human-Computer Interaction, Beijing, P.R. China, Springer Berlin / Heidelberg.

Thomas, B. H. and W. Piekarski (2002). "Glove based user interaction techniques for augmented reality in an outdoor environment." Virtual Reality **6**(3): 167-180.

Thomas, J. R. (1980). "Acquisition of motor skills: information processing differences between children and adults." Research Quarterly **51**(1): 158-173.

Thornton, S. (2002). Growing minds: An introduction to children's cognitive development, Palgrave.

Trankle, U. and D. Deutschmann (1991). "Factors influencing speed and precision of cursor positioning using a mouse." Ergonomics **34**(2): 161-174.

Unity Technologies. (2016). "Unity - Game Engine." Retrieved August 1, 2016, from https://unity3d.com/.

Vanacken, L., T. Grossman and K. Coninx (2009). "Multimodal selection techniques for dense and occluded 3d virtual environments." International Journal of Human-Computer Studies **67**(3): 237-255.

Vasilyeva, M. and S. F. Lourenco (2010). "Spatial development." The Handbook of Life-Span Development.

Vasta, R., K. G. Regan and J. Kerley (1980). "Sex differences in pattern copying: Spatial cues or motor skills?" Child Development: 932-934.

Volman, M., E. van Eck, I. Heemskerk and E. Kuiper (2005). "New technologies, new differences. Gender and ethnic differences in pupils' use of ICT in primary and secondary education." Computers & Education **45**(1): 35-55.

Voyer, D., A. Postma, B. Brake and J. Imperato-McGinley (2007). "Gender differences in object location memory: A meta-analysis." Psychonomic Bulletin & Review **14**(1): 23-38.

White, S., D. Feng and S. Feiner (2009). Interaction and presentation techniques for shake menus in tangible augmented reality. Mixed and Augmented Reality, 2009. ISMAR 2009. 8th IEEE International Symposium on, IEEE.

Woodcock, R. W. (1990). "Theoretical foundations of the WJ-R measures of cognitive ability." Journal of Psychoeducational Assessment **8**(3): 231-258.

Wyeth, P. a. P., H.C. (2003). "Using developmental theories to inform the design of technology for children." Proceedings of the 2003 conference on Interaction design and children: 93-100.

Xu, Y., S. Mendenhall, V. Ha, P. Tillery and J. Cohen (2012). Herding nerds on your table: NerdHerder, a mobile augmented reality game. Proceedings of the 2012 ACM annual conference extended abstracts on Human Factors in Computing Systems Extended Abstracts, ACM.

Yan, J. H., J. R. Thomas, G. E. Stelmach and K. T. Thomas (2000). "Developmental features of rapid aiming arm movements across the lifespan." Journal of Motor Behavior **32**: 121-140.

Zhou, Z., A. D. Cheok, J. H. Pan and Y. Li (2004). Magic Story Cube: an interactive tangible interface for storytelling. Proceedings of the 2004 ACM SIGCHI International Conference on Advances in computer entertainment technology, ACM.

Zhou, Z. Y., A. D. Cheok, J. Tedjokusumo and G. S. Omer (2008). "wIzQubesTM–A novel tangible interface for interactive storytelling in mixed reality." International Journal of Virtual Reality **7**: 9-15.